DOCUMENT RESUME

ED 039 321 08 VT 009 998

AUTHOR Loudermilk, Kenneth M.; DiMinico, Gerald
TITLE Instruments for Vocational Guidance, Selection, and Placement: A Review and Synthesis of Research in Idaho.
INSTITUTION Idaho Occupational Research Coordinating Unit, Moscow.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
BUREAU NO BR-6-3014
PUB DATE Mar 69
GRANT OEG-4-7-063014-1590
NOTE 118p.

EDRS PRICE EDRS Price MF-$0.50 HC-$6.00
DESCRIPTORS Abstracts, Aptitude Tests, Doctoral Theses, *Job Placement, Masters Theses, *Occupational Guidance, Research Reviews (Publications), *Selection, Test Interpretation, Test Reliability, *Tests, Test Selection, Test Validity, *Vocational Counseling
IDENTIFIERS GATB, General Aptitude Test Battery, *Idaho

ABSTRACT

The development and use of instruments for vocational guidance, selection, and placement within the State of Idaho are reviewed. Vocational guidance is defined as assisting the individual to understand himself, the world of work, and career choice. Selection and placement are described as activities characteristically used by educational institutions and business organizations in deciding who will be accepted and what roles and treatments would be expected of those who are accepted. Specific attention is given to research in which the General Aptitude Test Battery (GATB) was used to predict success in vocational training or work performance. Separate chapters are devoted to research studies with appraisal instruments completed in Idaho as well as 31 studies done elsewhere in the nation. Because more than three-fourths of the studies were completed as individual graduate research papers or theses, research results were generally not comparable or additive from one study to another. An extensive bibliography is included. An earlier study "A Survey of Literature Related to Selected Nonprofessional Occupations" is available as ED 013 915. (CH)

ED039321

The State Occupational Research Unit
College of Education
University of Idaho
Moscow, Idaho

INSTRUMENTS FOR VOCATIONAL GUIDANCE, SELECTION, AND PLACEMENT: A REVIEW AND SYNTHESIS OF RESEARCH IN IDAHO

Kenneth M. Loudermilk, Director
Gerald DiMinico, Research Fellow

The work reported herein was performed pursuant to Grant Number OEG-4-7-063014-1590 from the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

VT009998

March, 1969

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FOREWORD

This report is the second in a series which deals with the selection and use of several types of individual appraisal instruments in Idaho's system of vocational-technical education. The first report, published in May, 1966 by the State Occupational Research Unit, was entitled "A Survey of Literature Related to Selected Nonprofessional Occupations." The professional literature was surveyed for studies related to success and/or membership in 28 occupations for which vocational training programs existed in Idaho. The survey revealed that about 10 broad types of appraisal instruments profitably could be incorporated into an experimental test battery.

The present study focused more on empirical research studies. An effort was made to locate all research studies in Idaho which pertained to the construction and use of appraisal instruments for vocational guidance, selection, and placement. Selected research studies completed or in progress elsewhere in the nation also were cited and discussed which had relevance for future developments in Idaho.

This second report is intended as another interim report, as much more research remains to be completed in Idaho. A third report might present the results of a survey and critical analysis of tests and other appraisal instruments which should compose an experimental test battery. Immediately following this, a test validation study involving all of the area vocational schools should begin. Additional reports would be issued from time to time as the project developed and as research data accumulated. Whether or not any additional reports actually are issued depends upon how much research activity can be generated, and the amount of research money that is made available.

The writer is indebted to several persons throughout the state whose assistance in various phases of this project eased the burden of library research, and who made valuable suggestions during informal conversations which influenced the writer's thinking. The following five individuals assisted in locating pertinent research reports: Mr. Bill Gard, Vocational Counselor, Idaho State University, Pocatello; Dr. William H. Hedley, Associate Dean of Students, College of Idaho, Caldwell; Dr. Dwight Kindschy, Head, Department of Agricultural Education, University of Idaho, Moscow; Dr. Richard A. Merriman, Regional Director, Idaho Continuing Education, Idaho State University, Pocatello; and Mr. Hugh Nelson, Counseling and Testing Supervisor, Department of Employment, State of Idaho, Boise. During the summer of 1966, two counselors assisted in the collection of data for test validation studies at their respective area vocational schools: Mr. Len Brenchley, Vocational Counselor, Idaho State University, Pocatello; and Mr. Charles F. Williamson, formerly Vocational Counselor, Boise State College, and presently Counseling Psychologist, Veterans Administration, Boise. Finally, the writer wishes to thank Dr. Oscar E. Kjos, Coordinator of Vocational Teacher Education, University of Idaho, who several times acted as a "sounding board" and as a stimulator and contributor of important research ideas and interpretations. Needless to say, none of these gentlemen can be held responsible for the writer's selection and use of their suggestions and contributions, and for the conclusions and recommendations stated in this report.

K.M.L.

TABLE OF CONTENTS

CHAPTER PAGE

I. INTRODUCTION, PROBLEM, AND RESEARCH PROCEDURE 1
  INTRODUCTION 1
    A Point of View 1
    Recent Developments in Career Appraisal 3
  PROBLEM 5
    Different Emphases in Appraisal 5
    Appraisal Instruments--Basic Functions 5
    Research in Idaho 7
  RESEARCH PROCEDURE 8
    Research Methods 8
    Report Organization 9
II. THE GENERAL APTITUDE TEST BATTERY AND RELATED RESEARCH . . 10
  THE GATB--TECHNICAL ASPECTS 11
    Aptitudes Measured 11
    Reliability 11
    Validity 12
    Test-Retest Variability 12
  GATB USE IN TRAINING PROGRAMS 15
    Vocational Training Programs in Idaho 15
    Validity Data for Specific Occupations 18
    Variations in Predictive Studies 21
    Predictive Studies for Training Groups 22
  DISCUSSION, SUMMARY, AND CONCLUSIONS 22
    Discussion 22
    Summary 25
    Conclusions 25
III. VALIDATION STUDIES CONDUCTED BY THE STATE OCCUPATIONAL RESEARCH UNIT 27
  RESULTS FOR BOISE STATE COLLEGE 28
    The Correlation Studies 28
    The Analysis of Variance Studies 31
  RESULTS FOR IDAHO STATE UNIVERSITY 32
    The Correlation Studies 32

TABLE OF CONTENTS (CONT'D)

CHAPTER PAGE

The Analysis of Variance Studies 35

DISCUSSION, SUMMARY, AND CONCLUSIONS 36

Discussion 36

Summary 37

Conclusions 38

IV. A REVIEW OF RESEARCH IN IDAHO 39

IDAHO RESEARCH STUDIES 40

The Ayarra Study (1967) 40

The Bertrand Study (1966) 41

The Brenna Study (1961) 42

The Brown Study (1968) 43

The Burke Study (1967) 44

The Cain Study (1969) 46

The Chatburn Study (1956) 47

The Faulks Study (1967) 48

The Fifield-Jones Study (1964) 48

The Fifield-Watson Study (1967; Watson, 1967) 49

The Flores Study (1966) 50

The Gardner Study (1957) 51

The Gentry Study (1967) 52

The Holmquist Study (1967) 53

The Johnston Study (1967) 53

The Kiler Study (1967) 54

The Lee Study (1958) 57

The Loudermilk Study (1964, 1966) 57

The McEwen Study (1962) 59

The McHugh Study (1967) 60

The McRae Study (1968) 61

The Meyerhoeffer Study (1967) 61

The Nybroten Study (1964) 63

The Olson Study (1966) 63

The Rutledge Study (1967) 64

The Scharf Study (1966) 65

TABLE OF CONTENTS (CONT'D)

CHAPTER PAGE

The Sorensen Study (1967) 66
The Stewart Study (1968) 67
The Williamson Study (1964) 68
The Witty Study (1962) 69
The Young Study (1965) 70
DISCUSSION, SUMMARY, AND CONCLUSIONS 70
Discussion 70
Summary 72
Conclusions 72
V. NATIONAL DEVELOPMENTS IN VOCATIONAL APPRAISAL 76
HISTORICAL PERSPECTIVE 76
RECENT STUDIES AND REVIEWS 77
CAREER CHOICE AND DEVELOPMENT THEORY 79
COMPUTERIZED INFORMATION SYSTEMS 81
DISCUSSION, SUMMARY, AND CONCLUSIONS 82
Discussion 82
Summary 83
Conclusions 83
VI. SYNTHESIS, SUMMARY, AND CONCLUSIONS 84
SYNTHESIS 84
To Test Or Not To Test 84
Appraisals Without Tests 86
Appraisals With Tests 87
Outlook In Idaho 89
SUMMARY 90
CONCLUSIONS 91
BIBLIOGRAPHY 93
APPENDIX A 99
APPENDIX B 104

LIST OF TABLES AND FIGURES

TABLE PAGE

I. Study Programs in Idaho Area Vocational-Technical Schools and Specific Occupations With Norms From The GATB 16

II. Validity Data for Specific Occupations From The GATB . . . 19

III. Phi Coefficients for Training Groups Compared With The Highest Phi Coefficients Obtained in Earlier Studies With The GATB 23

IV. Correlations Between Nelson-Denny Reading Scores and Cumulative Grade Point Averages (Boise State College) . . . 29

V. Correlations Between GATB Scores and Cumulative Grade Point Averages (Boise State College) 30

VI. Correlations Between GATB Scores and Cumulative Grade Point Averages (Idaho State University) 34

FIGURE PAGE

1 Frequency Distribution of Phi Coefficients (N = 425) For Specific Occupations From The GATB 13

CHAPTER I

INTRODUCTION, PROBLEM, AND RESEARCH PROCEDURE

INTRODUCTION

A Point of View

It seems desirable at the very beginning of this report to present a point of view which is fundamental to a consideration of the assessment of individuals for guidance, selection, and placement. The following quotation taken from the writings of Munsterberg, the father of applied psychology, seems to get to the heart of the matter.

> Like all technical sciences, applied psychology tells us what we ought to do if we want to reach certain ends; but we ought to realize at the threshold where the limits of such a technical science lie, as they are easily overlooked, with resulting confusion. We must understand that every technical science says only: you must make use of this means, if you wish to reach this or that particular end. But no technical science can decide within its limits whether the end itself is really a desirable one. The technical specialist knows how he ought to build a bridge or how he ought to pierce a tunnel, presupposing that the bridge or the tunnel is desired. But whether they are desirable or not is a question which does not concern the technical scientist, but which must be considered from economic or political or other points of view. Everywhere the engineer must know how to reach an end, and must leave it to others to settle whether the end is in itself desirable. Often the end may be a matter of course for every reasonable being. The extreme case is presented by the applied science of medicine, where the physician subordinates all his technique to the end of curing the patient. Yet if we are consistent we must acknowledge that all his medical knowledge can prescribe to him only that he proceed in a certain way if the long life of the patient is acknowledged as a desirable end. The application of anatomy, physiology, and pathology may just as well be used for the opposite end of killing a man. Whether it is wise to work toward long life, or whether it is better to kill people, is again a problem which lies outside the sphere of the applied sciences. Ethics or social philosophy or religion have to solve these preliminary questions. (Munsterberg, 1913, pp. 17-18)

This quotation is appropriate here for at least two reasons. First, few persons, be they laymen or top professionals in the field, are unaware of the heated controversies that have arisen over the use of tests and other assessment instruments. Tests and other similar instruments have been discussed as if they constituted some decision-making force, whereas in reality they are merely tools (or means) used to achieve certain ends decided upon by people. It is of more than passing interest that Munsterberg saw this distinction more than half a century ago, and related it specifically to the field of applied psychology which was just developing.

The second reason for citing Munsterberg is even more important, and bears directly on the contents of this report. At the risk of being a bit facetious, the point may be illustrated with a brief story. One man said to another, "How is your wife?" The other replied, "Compared to what?" This type of question is implicit in any criticism of the use of tests. The writers almost never consider what might happen in the complete absence of the testing program they criticize. Actual conditions in such a case--and, indeed, conditions present at the time but not affected directly by the testing program--might reflect even more chaos and poor decisions than those resulting from the misuse of tests. None of the assessment instruments developed to date are perfect, and no responsible individual has ever said that they were. But decisions are a necessary part of life and, other things equal, those concerning people should be improved by the use of instruments which improve our understanding, prediction, and to some extent the control of human behavior. This is perhaps the main reason for the development of tests and other similar devices. We do well to examine and evaluate them critically, but in all fairness we need to compare them with any and all other means which may be used in our work with people.

One other notable individual deserves mention, who was a contemporary of Munsterberg early in this century. This is Parsons, the father of vocational guidance. Parsons recognized the value of individual assessment in guiding the individual vocationally, but his methods and tools reflect the psychological developments of his time.

> Rapidity can be judged by testing the swiftness of reading and writing and walking, and if a psychologic laboratory or some psychologic apparatus is available, it is easy to apply much more accurate tests through the phenomena of reaction-time, association-time, etc., than are readily available without scientific apparatus.
>
> If the applicant's head is largely developed behind the ears, with big neck, low forehead, and small upper head, he is probably of an animal type, and if the other symptoms coincide he should be dealt with on that basis.
>
> If the voice is harsh, or unpleasant, or lacking in vitality, I generally give the youth a lecture on the value of voice culture and the use of clear, sweet, well-modulated tones in conversation.
>
> If the face is blank and expressionless, a talk about the economic value of the smile is in order.
>
> If the handshake is listless or wet, clammy or too forceful, it is well to call the young man's attention to his defects in this respect. So if the manners are in any way objectionable or undeveloped, the boy should be frankly but kindly told and urged to correct his errors. (Parsons, 1909, pp. 22-23)

The writer of this report purposely lifted this excerpt out of context to illustrate some of the earliest efforts to use individual assessments in vocational guidance. Parsons generally was on the right track as regards the major tasks to be completed; his means were sometimes naive by present-day standards. We need, however, to consider some modern developments lest we assume that we have "arrived" in the field of vocational guidance and the attendant use of assessment instruments.

Recent Developments in Career Appraisal

That a vast amount of time and effort has been expended in the development of tests is evident from a recent book edited by Buros (1961). This book contains entries for nearly three thousand tests, and the number has grown since. It is a little disconcerting, therefore, to see a statement like the following from a recognized authority on the use of tests for vocational appraisal.

> When we ask which of these hundreds of tests should be used in counseling, we are usually posing the question: Which of these tests will best predict the future vocational adjustment of a client? By vocational adjustment we generally mean the client's success and satisfaction on the job after he has entered the world of work.
>
> The sad but true conclusion which we must draw is that most of our assessment instruments have little or no predictive validity in forecasting these two major criteria of vocational adjustment. (Crites, 1968, p. 23)

An important development in the last 20 years has been the appearance of a number of theories of vocational choice and development. These theories were developed at least partially in response to the general dissatisfaction with more conventional forms of guidance testing and appraisal. An important objective in the development of several of the career choice and development theories was to portray the choice of a career as a process which develops over time, rather than something that can be pinpointed in time. At least one theorist, however, seems to have his doubts as to whether all of these efforts are really worth the candle.

> . . . we must either make more efficient predictions on the basis of theory or abandon our efforts to construct theories. Despite several decades of research, the most efficient way to predict vocational choice is simply to ask the person what he wants to be; our best devices do not exceed the predictive value of that method. (Holland, 1966, p. 91)

Magazine articles which reach conclusions similar to those expressed by Holland and Crites, and a few books, have appeared for popular consumption. Most persons at all familiar with the testing field can name several, and it does not seem necessary to cite examples in this report. If we accept such conclusions, then the obvious course is to do away with all testing and similar assessments. Depending upon how we interpret the field, this also might require us to dispense with the assignment of school grades, and the awarding of certificates and diplomas. Decisions concerning people would still need to be made, however, and it is hard to imagine how these could be executed without _some_ kind of individual assessment.

The writer has adopted the view that there _are_ instruments available that have potential value for guidance, selection, and placement, despite some obvious imperfections and a number of abuses which have occurred. Crites, cited above, discussed in the same article his work on a new instrument called the Vocational Development Inventory which is expected to improve upon certain shortcomings of present instruments. Moreover, he recently served as coauthor with Super of a book (Super & Crites, 1962) which presents comprehensive reviews and evaluations of psychological tests for vocational appraisal. Holland, also cited previously, has developed his own theory of vocational choice and has completed a respectable amount of research to test the theory. These men apparently have enough faith in their endeavors to continue their work, even though they readily admit serious shortcomings in our efforts to understand and predict vocational behavior.

It seems necessary also to reflect briefly on the competencies of those persons who use tests. Even the simplest ones need to be administered and interpreted by someone who has been trained. We were told this more than 40 years ago by Lewis M. Terman in his introduction to Hull's _Aptitude Testing_.

> Of course it is possible for one who has no scientific understanding of test methods to learn the simple procedures of test administration and test scoring, or even to carry on a certain amount of routine manipulation of test data, but there is a growing recognition of the dangers involved in this kind of work. (Terman, in Hull, 1928, p. xiii)

Hull himself stated it as follows:

> In a word, aptitude testing, like medicine and engineering, is ceasing to be a job for amateurs and is becoming the work of technically trained professionals. (Hull, 1928, p. v)

Moreover, the father of vocational guidance closed his excellent treatise by stating that career guidance should be administered by ". . . experts trained as carefully for the work as men are trained to-day for medicine and the law." (Parsons, 1909, p. 165) If we had paid more attention to these admonitions over the last 60 years, it is almost certain that the present status of the field would be vastly different. A basic premise of the writer thus may be stated, and will be restated and illustrated elsewhere in this report: If trained personnel and facilities are not available to do a job, it should not be attempted at all.

PROBLEM

The basic problem which led to the preparation of this report is reflected in the title: "Instruments for Vocational Guidance, Selection, and Placement: A Review and Synthesis of Research in Idaho." Actually, three different lines of thought can be identified, and deserve an extended treatment in this section.

Different Emphases in Appraisal

Individual appraisal by means of psychological tests and similar instruments is a fundamental aspect of the three closely related--but not identical--functions of vocational guidance, selection, and placement. These terms have been used frequently in the literature. No definitions will be given, but a brief discussion is in order regarding the different emphases of each.

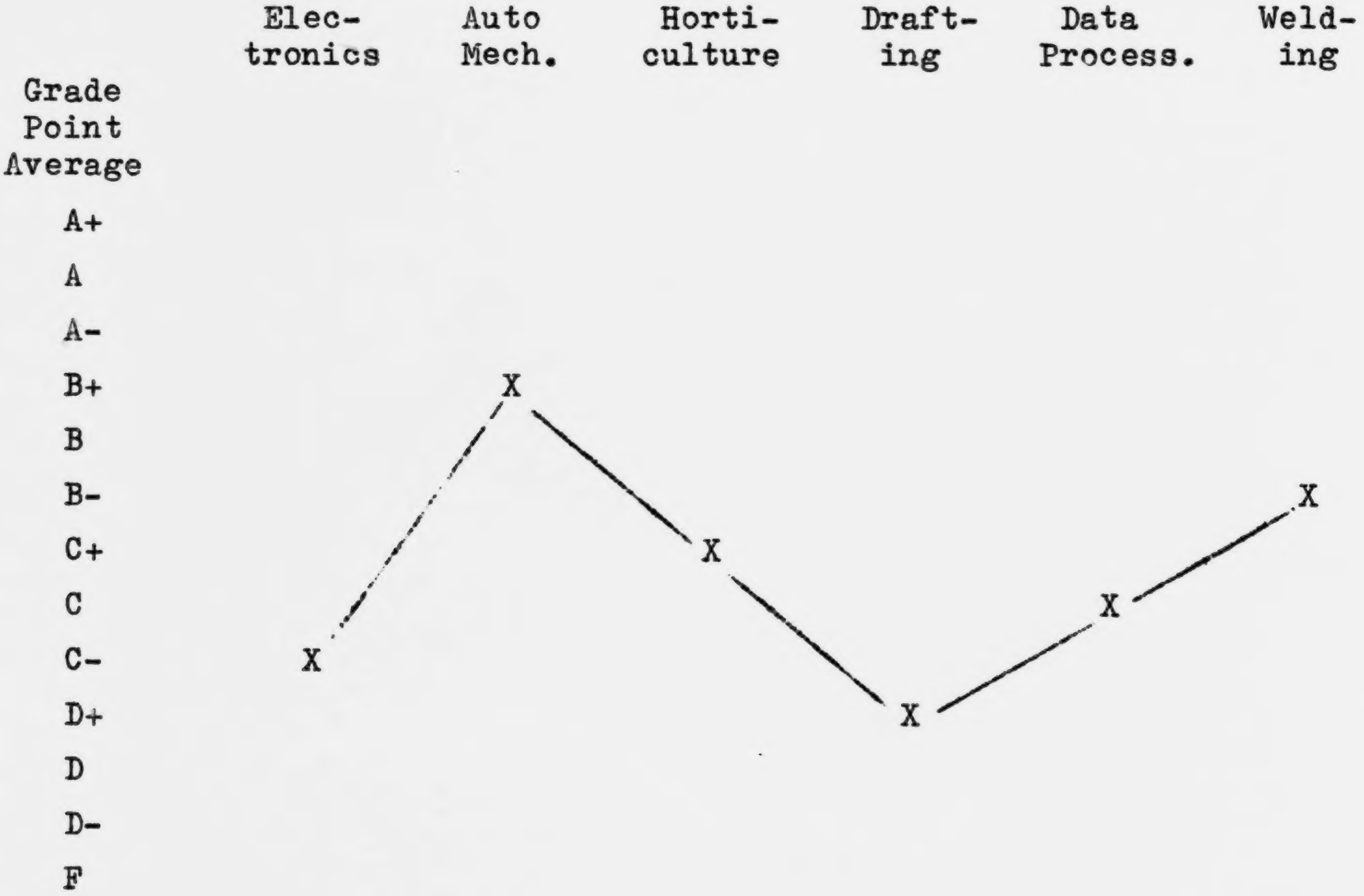
Vocational guidance generally is concerned with the individual as he views the world of work, including the many opportunities for education and training which lead into different types of employment. A basic objective is to help the individual to understand himself, the opportunities available to him, and to relate himself to the latter in a way that results in maximum benefits to both himself and society. The emphasis tends to shift away from the individual when selection and placement are considered. Decisions must be made by organizations such as schools and businesses regarding who will be accepted and what role they will play following acceptance.

Super and Crites (1962) and Cronbach and Gleser (1957) discuss all of the foregoing concepts at greater length. The point to be emphasized here is that the three terms have a different emphasis. Confusion results because frequently the same instruments used in vocational guidance also are used in the selection and placement of individuals. The three uses can conflict. For example, an individual counselee may decide upon a given field, only to find that he is not selected for training after applying at several schools. Or he might be selected by a school which, for various reasons, is unable to place him in the training program of his choice. Basically, both for individuals and organizations, decisions--or choices among alternatives--must be made. Even though conflicts are a possibility, this writer believes there also is considerable room for decisions which can benefit all concerned.

Appraisal Instruments--Basic Functions

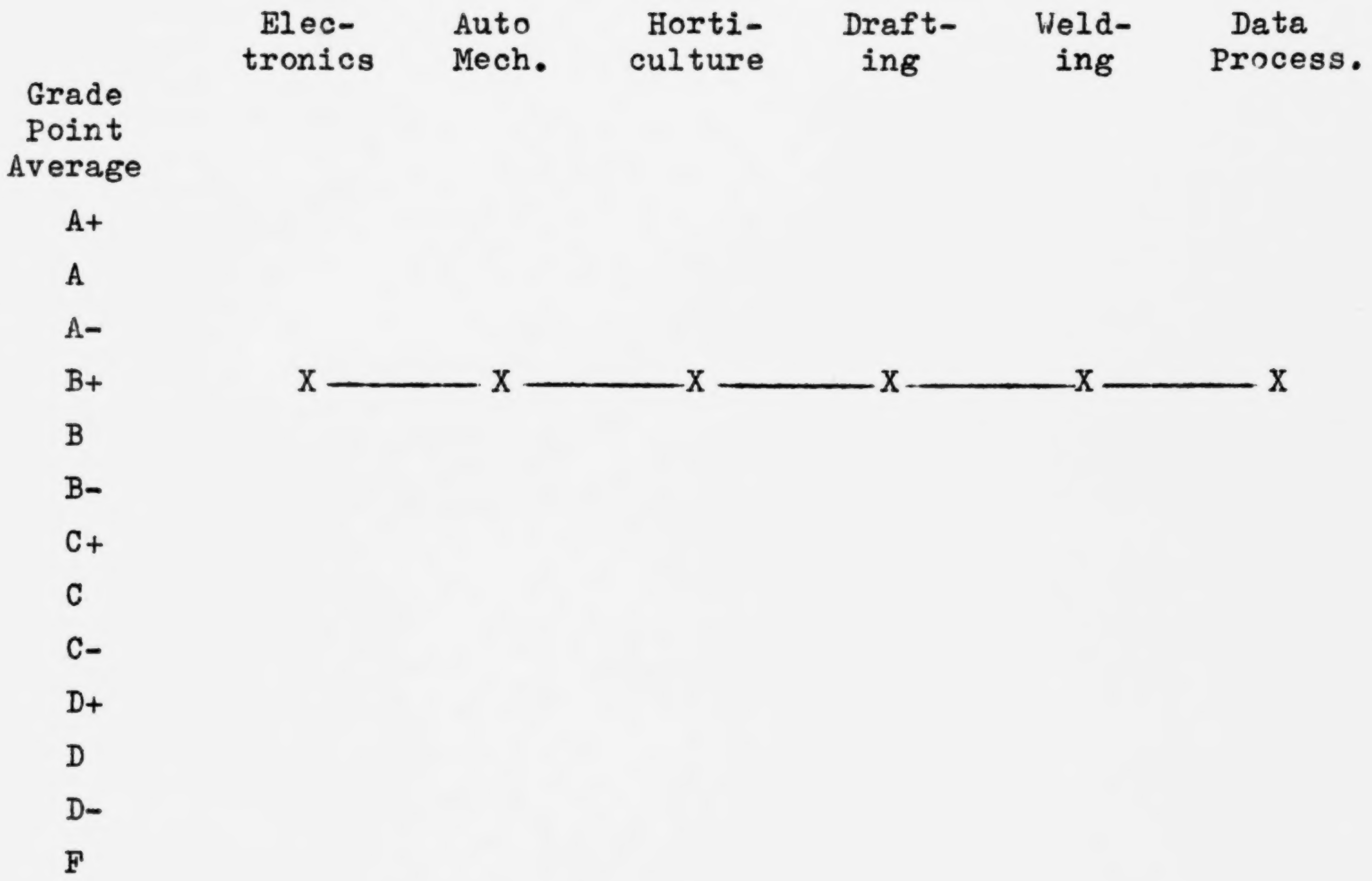
Two fundamental uses of appraisal instruments may be identified in the understanding and prediction of vocational behavior. Early in the 1950 decade a symposium at Harvard University was concerned with the newly developed multiple discriminant function (Tiedeman, *et al.*, 1951). The method also was compared with multiple regression analysis. The discussion which follows is based on the series of articles which resulted from the Harvard symposium as well as research publications which have appeared in the intervening years.

The most familiar use of appraisal instruments is the insertion of a student's scores in a regression formula which yields a predicted grade point average or some other criterion measure of academic success. Predictions for a hypothetical student might yield the following.



These predictions should be quite useful in helping a student decide among the several training alternatives. They also could be used by an admissions officer for selection and placement. We might assume, for sake of simplicity, that this student was free to choose, and that he selected auto mechanics, the program in which he appears to have the best chance of academic success.

The second use of appraisal instruments is less familiar, and may be illustrated by predictions of academic success for another hypothetical student (see illustration on the next page). This student has equally good academic potential for any and all of the six vocational programs. Indeed, a glance at the results immediately poses another question--a choice among alternatives. Assuming once again that the student is free to choose, his question might be as follows: "Which group of students do I resemble most in respect to interests, background traits, and other attributes?" Many students can (and do) rule out a number of programs which do not appeal to them, or which are



unsuitable for other reasons. It is possible, however, to develop a testing program so that information is available both for academic predictions within groups and for predictions concerning group membership.

The two uses of test results discussed above frequently are confused. Academic predictions (within groups) generally are best when various cognitive measures such as intelligence, aptitude, and achievement are included in an assessment battery. Interest measures, on the other hand, are relatively poor predictors of success within groups, as they measure traits in the affective domain. Conversely, choices between groups, or group membership, tends to be predicted best when non-cognitive measures such as interests and personality are used. Several studies during recent years (e.g. Berdie, 1955; Strong, 1955; Thorndike & Hagen, 1959) have been concerned more or less directly with the two basic uses of tests discussed here. Chapter V of this report discusses further the development of prediction batteries, and cites a number of studies which have important implications for similar studies in Idaho and elsewhere.

Research in Idaho

The present report is the second in a series of reports planned by the State Occupational Research Unit concerned with the use of instruments for vocational appraisal. The first report (State Occupational Research Unit, 1966) contained the results of a literature survey of studies relating to success and/or membership in 28 occupations for which training was available in Idaho. The present study focused on research studies completed in Idaho. Furthermore, appraisal instruments were evaluated which have been used in the state for guidance, selection, and placement in the vocational-technical area. At least one more study is needed, which would be a survey and critical evaluation of specific tests and related instruments which could be selected or developed for an assessment battery in Idaho's vocational-technical education system. The proposed study also would include administration of the battery thus selected, and subsequent research analyses. No specific plans have been made, however, for this third study in the series.

The present report is timely for at least two reasons. First, vocational guidance and research studies relating thereto have received increasing attention in recent years. The annual conventions of the American Vocational Association, for example, recently have included several programs in vocational guidance and related research. Second, the Center for Vocational and Technical Education, The Ohio State University, Columbus, Ohio, has contracted with several individuals to prepare Review and Synthesis Reports on important topics in the vocational-technical field. These reports apparently have made a significant contribution to the research program. It seemed that a similar effort in Idaho would be desirable and necessary to encourage further developments and improvements in the use of instruments for vocational appraisal. There is, of course, no direct connection between the present report and the series sponsored by the Ohio State Center.

RESEARCH PROCEDURE

This final section of the chapter is concerned with two main topics. The first is a brief account of the research methods used throughout the study. The second topic pertains to the organization of the remaining chapters.

Research Methods

The research methods used will become clearer as the reader progresses through the remainder of this report. At this point five research methods or strategies (in a generic sense) may be identified. First, the writer reviewed carefully the technical manual and additional research publications for the General Aptitude Test Battery. This battery has had extensive use in Idaho, and the manual is one of the best available in the testing field. Second, the writer engaged directly in research studies concerned with predictions of academic or vocational success and membership in certain defined groups. Standard statistical methods such as simple correlation, multiple correlation, analysis of variance, multivariate analysis of variance, and multiple discriminant analysis were employed. Third, all research studies in

Idaho which could be identified and obtained were reviewed and summaries of each were prepared. Fourth, the writer has attended a number of regional and national conventions and seminars where references and materials were available. A number of those deemed most important were selected for inclusion in this report. Finally, an effort was made to synthesize all of the material reviewed or otherwise developed for this report, and conclusions and recommendations were stated relative to future developments in Idaho.

Report Organization

The remaining five chapters contain the specific outcomes of the general research methods described in the preceding paragraph. Chapter II presents a critical examination of the General Aptitude Test Battery and its use in Idaho. Chapter III contains the results of empirical studies completed by the writer--mainly with the General Aptitude Test Battery. Chapter IV is a review of all research studies that could be located which pertain to the use of assessment instruments in Idaho's vocational-technical education system, or closely related areas of activity. Chapter V contains brief presentations of a few significant studies and information sources in the nation which have important implications for future developments in Idaho. Chapter VI completes the report with a synthesis of the contents of earlier chapters, a summary, and conclusions. Chapters II through V each contain a final section entitled Discussion, Summary, and Conclusions. This material provides a quick overview of the contents of each chapter, and facilitated the writing of the final chapter.

CHAPTER II

THE GENERAL APTITUDE TEST BATTERY AND RELATED RESEARCH

The General Aptitude Test Battery (GATB) is probably the best known multi-factor battery in the vocational area. The GATB was constructed by the U.S. Employment Service and has a long history of development and research. The technical manual (U.S. Department of Labor, 1967a) is an exemplary contribution to the mental measurement field. The GATB generally has received favorable reviews and evaluative comments (e.g. Anastasi, 1968; Super, 1957; Super & Crites, 1962). There are, however, a number of areas in which improvements are needed, as evidenced especially by reviews in two recent Mental Measurements Yearbooks (Buros, 1959, pp. 692-700; 1965, pp. 1021-1029).

The GATB is discussed and evaluated in this report because it has had extensive use in Idaho in the Department of Employment, State of Idaho, and in the area vocational schools. There seem to be two main reasons which account for the latter use. First, the extensive validation studies have resulted in a number of specific aptitude batteries for occupations which logically can be matched with several of the vocational training programs offered in the area schools. Second, test materials, personnel to administer and score the tests, and additional related services have been available through the Department of Employment free of charge. Only recently has the Department of Employment asked the area schools to assume some of this responsibility, such as test administration.

Such extensive use calls for a critical look at the research literature accumulated for the GATB. Moreover, a few people throughout Idaho have questioned whether this battery is really effective as a guidance, selection, and placement tool. Research results for the GATB are available in large quantities, both in the technical manual and similar reports, and in psychological journals. The purposes of this report, however, call for a selective accumulation of research results which pertain directly to conditions existing in Idaho. Accordingly, it is expected that readers not satisfied with the selection or treatment here will avail themselves of the more extensive literature.

Another basic premise adopted by the writer, and adhered to throughout the remainder of this report, needs to be stated. The reader thus will know the basis upon which a number of conclusions are drawn. The correlation coefficient commonly is used to indicate the degree of relationship between variables, and gives information regarding the precision with which certain behaviors or attributes may be predicted. The choice of a given magnitude for a correlation is arbitrary, depending upon a number of circumstances. The writer will bypass further discussion and merely state that a correlation (either single or multiple) of .45 should be obtained before one or more tests can be deemed suitable as a basis for guidance, selection, or placement. This size of correlation is not without some logic and precedent, as Hull (1928, p. 275) suggested it more than 40 years ago as the minimum for work with individuals. Moreover, it is a level which can be attained in many well designed validation studies. The writer also decided that no decisions should be based on a test appraisal with a correlation of .30 or less. Another interpretation of the above correlations is their index of forecasting efficiency (Garrett, 1953, pp. 175-176), or the percentage predictions are improved over a "mere guess." A correlation of .45 represents an improvement over guesswork of about 11 per cent, and a correlation of .30 about 5 per cent. The writer believes that any appraisal, to be worthwhile, should at least be equivalent to odds of 1 in 10; odds of 1 in 20 do not seem worth the effort, and leave too much chance for error.

THE GATB--TECHNICAL ASPECTS

Technical considerations like reliability and validity are standard for most appraisal instruments. These and other aspects of the GATB, and research having important implications for its use in Idaho, are presented under the following headings.

Aptitudes Measured

The GATB provides nine measures derived from factor analyses of more than 50 tests early in the 1940 decade. Definitions of the aptitudes are available from other sources (U.S. Department of Labor, 1967a, p. 15). Only the name of each measure, and its letter symbol, are presented here: Intelligence or general learning ability (G), verbal aptitude (V), numerical aptitude (N), spatial aptitude (S), form perception (P), clerical perception (Q), motor coordination (K), finger dexterity (F), and manual dexterity (M). The battery can be conveniently administered in slightly over two hours. Aptitudes F and M require special apparatus, thus causing some inconvenience and additional administration time when large groups are tested. Scoring is by hand or by machine, and yields a profile of the nine aptitude scores. Each score is scaled to a mean of 100 and a standard deviation of 20, making possible direct comparisons between scores. Interpretation of the individual's aptitude profile relative to occupational groups is accomplished mainly through the use of the multiple cut-off method.

Reliability

A perusal of the manual section on reliability (U.S. Department of Labor, 1967a, pp. 193-215) reveals that the GATB yields fairly consistent results. Several studies report coefficients of stability (test-retest) and coefficients of equivalence (alternate forms) which range between the .80 and .90 levels. A few are .90 or higher, whereas a few coefficients, especially for the psychomotor tests (aptitudes F and M) are below the .80 level. Reliabilities of this magnitude, with the possible exception of the measures of aptitudes F and M, seem sufficiently high for general use.

Validity

Whereas reliability and certain other technical aspects are important, the real value of an appraisal instrument to understand and predict vocational behavior tends to focus upon validity. The GATB has been studied extensively in a number of settings and with a wide variety of groups. It probably is unsurpassed in regard to the amount of validation research which has been, and continues to be, accumulated. It is imperative, therefore, that this wealth of research information be evaluated carefully as it relates to the requirements of local programs in Idaho.

A good place to begin is Table 9-1 in the technical manual (U. S. Department of Labor, 1967a, pp. 60-81). Specific batteries composed of at least two and as many as four of the nine aptitudes measured by the GATB are selected in research studies as norms for specific occupations. The relationship between the specific battery thus selected and a criterion such as school grades or supervisory ratings is reported as a phi coefficient. The writer tallied all of the phi coefficients reported in Table 9-1, and the results are presented in Figure 1 on the page following. Table 9-1 has 385 numbered entries, whereas Figure 1 is based on 425 phi coefficients. The latter number is larger because several entries have more than one validation study.

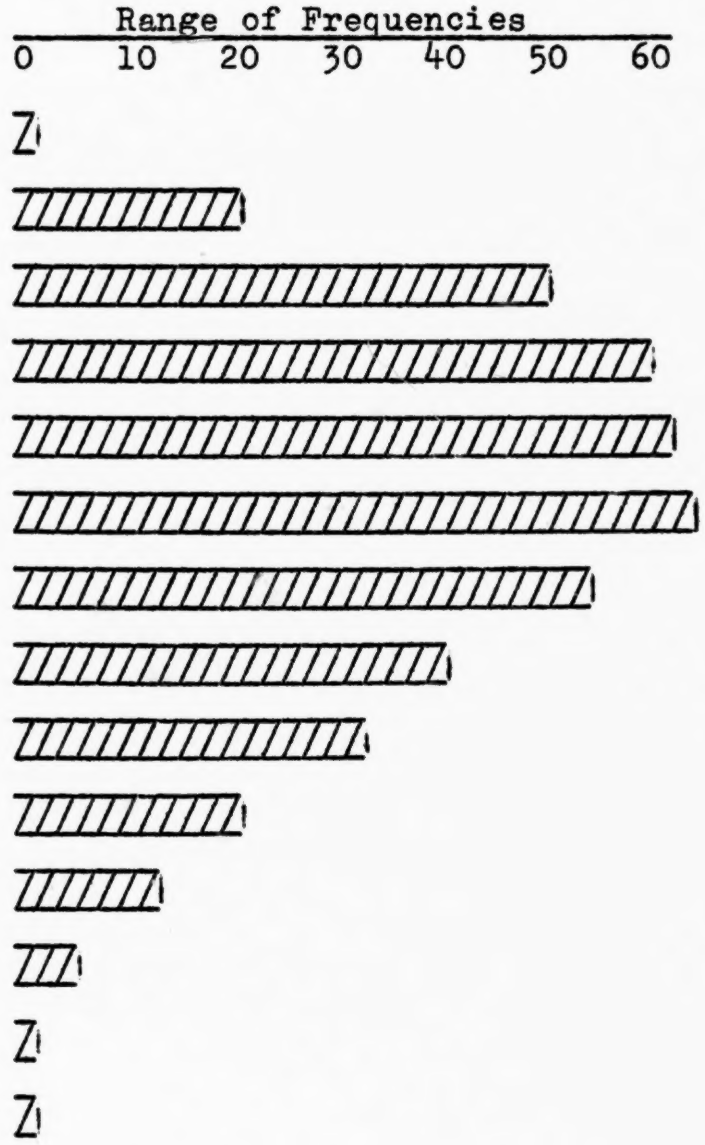
While this many separate studies is a remarkable achievement, an examination of Figure 1 reveals that the GATB should be used with considerable caution. The writer previously set .45 as the minimum correlation necessary for practical use. Figure 1 reveals that 258, or 61 per cent, of the phi coefficients fall below this level. Furthermore, a sizeable 74, or 17 per cent, are less than .30, a range the writer considers unsatisfactory for any instrument used as the sole basis for guidance, selection, and placement. On the other hand, several of the phi coefficients are quite impressive. It is of interest, however, that only two of them exceed .80, which bears out Hull's (1925) prediction of nearly 44 years ago of the maximum predictive efficiency we could expect from test batteries. A later section of this chapter reveals that even the high correlations reported do not automatically indicate that the GATB should be used without some local validation studies.

Test-Retest Variability

The amount of variability on a test from one administration to another is closely related to reliability, and both are discussed together in the same section of the GATB technical manual (U.S. Department of Labor, 1967a, pp. 193-215). Test-retest variability is, however, sufficiently important to receive separate consideration. The technical manual reports several studies which show a definite tendency for scores to increase on a subsequent testing due to practice

Figure 1

Frequency Distribution of Phi Coefficients (N = 425)
For Specific Occupations From The GATB



Phi Intervals	Frequency
.15-.19	2
.20-.24	21
.25-.29	51
.30-.34	60
.35-.39	61
.40-.44	63
.45-.49	55
.50-.54	39
.55-.59	33
.60-.64	19
.65-.69	13
.70-.74	5
.75-.79	1
.80-.84	2

effects. Standard errors of measurement range from around 5 score points for general learning ability to more than 10 points for the psychomotor tests. The latter figure represents more than half a standard deviation. Some studies report mean increases of 5 or more score points. As might be expected, the shorter test-retest intervals tend to produce the greatest increases from practice effects.

At least two important points must be considered in view of these test-retest differences. First, if these are mean differences, the score increases for certain individuals could be considerably larger. The writer has administered the GATB to hundreds of persons and has counseled with several concerning the results. A few cases which were retested scored so much higher the second time that the results did not appear to belong to the same person. One in particular brought his scores up from below average to nearly one standard deviation above the mean. Interpretations of such results, especially using the multiple cut-off method, is very difficult and the counselor probably should follow up with additional tests. This test-retest variability is not unique to the GATB. A relatively obscure article by van der Reis (1963) also reveals similar tendencies toward higher scores on a retest. A total of 1,688 persons had been tested and then retested on measures of mental alertness, mechanical comprehension, arithmetic, and abstract reasoning. The results revealed that between 76 and 87 per cent of the group obtained retest scores not significantly different from their initial results.

The above discussion leads logically to the second point, namely, the interpretation of test results. Testing programs tend to be organized so that a large number of students are tested at one time--and only once. Some students may claim that they weren't at their best that day, and could do better on a retest. According to the results presented in the preceding paragraph, the chances probably are better than 1 in 10 that they could do so. The GATB technical manual (U.S. Department of Labor, 1967a, p. 212) states that an individual's score theoretically could be expected to fluctuate about his "true" score, and that the fluctuation could be accounted for by the standard error of measurement. The study by van der Reis (1963) showed, however, that more than 1 person in 10 would differ on a retest by more than twice the standard error of measurement. The GATB studies failed to include similar individual analyses.

The foregoing results require some consideration when test validation studies are conducted, and especially when test results are applied to specific individuals. The study by van der Reis included some suggestions such as lengthening the time limits of tests, and permitting the testee to take parts of the battery again on which he scored markedly lower than on the other parts. Additional possibilities exist, one of which would be to allow individuals to take the same or similar forms of a test several times during their school career. Marked discrepancies then could be investigated. Studies could be completed in which moderators could be sought which would differentiate persons likely to score significantly higher on a retest from those who score essentially the same. The idea here is related to the recent work of Ghiselli (1963) on the prediction of individual error of measurement and error of prediction. Finally, validation study designs are needed which compare the predictive accuracy of an initial test result to that of a significantly higher retest result. It seems common to take the higher performance as indicative of the individual's "true" ability. Definitive studies are needed to determine if this is a justifiable procedure.

GATB USE IN TRAINING PROGRAMS

The GATB has been used for several years as a tool for guidance, selection, and placement in Idaho's area vocational schools. This section presents a list of the vocational programs available as of May, 1968 and research data from the GATB which relates directly to similar specific occupations.

Vocational Training Programs in Idaho

As of May, 1968, Idaho had five area vocational schools, and 36 separate programs were offered at one or more of the schools. This list of programs was compared with the contents of Table 9-1 in the technical manual (U.S. Department of Labor, 1967a, pp. 60-81) of the GATB. Specific occupations in Table 9-1 were paired with the vocational programs in all cases where the two titles seemed to have a logical relationship. The results are presented in Table I on the following two pages. The vocational programs are listed in the first (left-hand) column. The GATB table number is included for reference purposes, and is the number assigned to the entries in Table 9-1 of the technical manual along the left-hand margin. The title of the specific occupation from Table 9-1 is presented in the third (right-hand) column of Table I.

An examination of Table I reveals that a related occupational title from the GATB technical manual was found for 27, or 75 per cent, of the vocational programs. The reader might question some of the pairings, or might suggest a few additional entries in the GATB manual that could be paired with the remaining 9 programs for which the writer did not locate a matching occupation. The important point, however, is that the whole process is arbitrary. Nothing short of an extensive investigation would reveal how closely the training content of the Idaho vocational programs compares with the training received by the GATB validation groups or the kinds of work tasks they were called upon to perform. This consideration alone tends to cast some doubt on the use of the GATB as a guidance, selection, and placement tool in Idaho's vocational education system. Moreover, the fact that one-fourth of the vocational programs offered cannot be paired with a specific occupation from the GATB indicates that, if use of the GATB continues, local validation studies should be planned and conducted. The material to follow makes an even stronger case for research

TABLE I

STUDY PROGRAMS IN IDAHO AREA VOCATIONAL-TECHNICAL SCHOOLS AND SPECIFIC OCCUPATIONS WITH NORMS FROM THE GATB

Programs	Specific Occupations--GATB	
	Table Number	Occupation
Appliance Repair	113	Electrical-Appliance Serviceman
Auto Mechanics	19	Automobile Mechanic
Auto Service Station Specialist	20	Automobile-Service-Station Attendant
Auto Body and Fender	18	Automobile-Body Repairman
Aviation Mechanics (F.A.A. Approved)	2	Aircraft-and-Engine Mechanic
Business Machine Repair	240	Office-Machine Serviceman
Civil Engineering Technology	125	Engineering Aid II
Cosmetology	85	Cosmetologist
Data Processing	282	Programmer, Business
Dental Assistant	94	Dental Assistant
Diesel Mechanics	101	Diesel Mechanic
Distribution, Marketing, Mid-Management	210	Manager, Store I
Drafting and Design	109	Draftsman, Architectural; Mechanical; Structural
Electronics	121	Electronics Technician
Electronics Equipment Servicing	120	Electronics Mechanic
Food Services	145	Food Service Worker II
Industrial Electronics	114	Electrical Technology-Technical Institute Training
Industrial Mechanics	205	Maintenance Man, Factory or Mill

TABLE I (CONT'D)

STUDY PROGRAMS IN IDAHO AREA VOCATIONAL-TECHNICAL SCHOOLS AND SPECIFIC OCCUPATIONS WITH NORMS FROM THE GATB

Programs	Specific Occupations--GATB Table Number	Occupation
Instrumentation	176	Instrument Repairman I
Machine Shop	203	Machinist I
Office Occupations	37	Bookkeeper I
" "	331	Stenographer, Typist
Operating Room Assistant	341	Surgical Technician
Peace Officer Training	257	Patrolman
Practical Nursing	236	Nurse, Licensed Practical
Printing	274	Press Man Occupations, Selected
Upholstering	150	Furniture Upholsterer
Welding	380	Welder, Combination
Agri-Business		None
Child Care Services		None
Crop and Soil Technology		None
Clothing Services		None
Forestry Aide		None
Home Furnishing Services		None
Housekeeping Services (Institution and Home)		None
Ornamental Horticulture Technician		None
Visiting Homemakers		None

studies to be conducted within Idaho.

Validity Data for Specific Occupations

A wealth of data is presented in the technical manual of the GATB regarding the validity of norms developed for specific occupations. The writer selected four different kinds of data from Table 9-1 of the technical manual which seemed most important for the present report. The data selected are presented in Table II on the following two pages. The first (left-hand) column of Table II presents the title of the specific occupation--in the same order as in the third (right-hand) column of Table I. Moving to the right across Table II, data are presented on the sample, the number of cases, the criterion, and the phi coefficient obtained in each study. Each of the four sets of data are discussed in the paragraphs which follow.

The Samples for 42 separate studies (56 per cent larger than the 27 specific occupations because several have more than one study) represent mainly employees and students, but also applicants, trainees and apprentices. A count reveals that 21 of the studies, or exactly half, were based on student samples (including Machinist I which includes both students and employees). Ideally for guidance purposes, several different samples should be included in test validation. Use of the GATB for selection and placement, however, is not strongly indicated for vocational programs in Idaho because too many of the specific occupations listed in Table II have no studies involving students. Ghiselli (1955, 1966) found important differences in the predictive value of tests for training criteria as compared with proficiency criteria. Researchers in the U.S. Employment Service (U.S. Department of Labor, 1967b) also found differences in this regard sufficiently large to indicate the desirability of separate studies for each type of criterion.

The Number of Cases was not large enough in any of the studies included in Table II for the results to be generalized and used for selection and placement in a new setting. The writer, through a personal communication, learned that none of the studies entered in Table II were based on cases residing in Idaho. We are, therefore, entirely dependent upon predictive relationships found in studies in other sections of the country. Strong (1943, p. 649), more than a quarter of a century ago, told us that he found it necessary to use a sample of from 300 to 500 cases to develop stable scales for his interest inventory. More recently, Nunnally (1967, p. 165) recommended the same number of cases in multiple correlation studies in which there are as many as 10 variables. Only 9 of the 42 studies, or 21 per cent, in Table II have more than 100 cases, and only one has over 200 cases. Once again, local validation studies within Idaho are indicated before the GATB can be recommended for general use.

The Criteria in Table II are of many different kinds. Since space limitations required a number of abbreviations, they are written here: Supervisory Ratings, School Grades, Instructor's Ratings,

TABLE II

VALIDITY DATA FOR SPECIFIC OCCUPATIONS FROM THE GATB

Specific Occupations--GATB	Sample	N	Criterion	phi
Electrical-Appliance Serviceman	Employees	53	Superv. Ratings	.33
Automobile Mechanic	Students	50	School Grades	.40
Automobile-Service-Station Attendant	Employees	52	Superv. Ratings	.29
Automobile-Body Repairman	Students	56	Instr. Ratings	.55
Aircraft-and-Engine Mechanic	Students	75	Grade-Point Ave.	.46
Office-Machine Serviceman	Trainees	62	Superv. Ratings	.50
" " "	Employees	55	Superv. Ratings	.40
Engineering Aid II	Students	57	Grade-Point Ave.	.42
Cosmetologist	Employees	99	Instr. Ratings & School Grades	.38
Programmer, Business	Employees	102	Superv. Ratings	.35
" "	Employees	93	Superv. Ratings	.31
Dental Assistant	Students	53	Instr. Ratings	.53
" "	Students	85	Instr. Ratings	.22
" "	Students	121	Superv. Ratings	.66
" "	Employees	31	Superv. Ratings	.44
Diesel Mechanic	Trainees	55	Instr. Ratings	.33
Manager, Store I	Students	51	Grade-Point Ave.	.26
Draftsman, Architectural;	Employees	52	Superv. Ratings	.30
Mechanical; Structural	Students	93	Instr. Ratings & GPA	.23
Electronics Technician	Students	97	Grade-Point Ave.	.52
" "	Students	51	Grade-Point Ave.	.36
Electronics Mechanic	Employees	50	Superv. Ratings	.45
Food Service Worker II	Employees	100	Superv. Ratings	.20
Electrical Technology-Technical Institute Training	Students	63	Grade-Point Ave.	.42

TABLE II (CONT'D)

VALIDITY DATA FOR SPECIFIC OCCUPATIONS FROM THE GATB

Specific Occupations--GATB	Sample	N	Criterion	phi
Maintenance Man, Factory or Mill	Applicants	53	Superv. Ratings	.83
Instrument Repairman I	Trainees	65	Course Grades	.35
" "	Trainees	58	Instr. Ratings	.26
Machinist I	Employees & Students	111	Superv. Ratings & Instr. Ratings	.36
Bookkeeper I	Students	66	Instr. Ratings	.49
Stenographer, Typist	Students	130	Work Sample	.22
" "	Students	60	Work Sample	.47
" "	Students	50	Work Sample	.40
" "	Students	58	Work Sample	.31
Surgical Technician	Employees	50	Superv. Ratings	.53
" "	Employees	52	Superv. Ratings	.28
Patrolman	Employees	166	Superv. Ratings	.22
"	Employees	64	Superv. Ratings	.30
Nurse, Licensed Practical	Students	119	GPA & PNL	.27
" " "	Students	111	Grade-Point Ave.	.57
Press Man Occups., Selected	Applicants, Employees, Apprentices	293	Superv. Ratings	.37
Furniture Upholsterer	Employees	90	Superv. Ratings & Produc. Records	.45
Welder, Combination	Students	84	Instr. Ratings	.64

Grade-Point Average (GPA), Course Grades, Work Sample, Practical Nurse License (PNL), and Production Records. Idaho's area vocational schools seem to place a major emphasis on grades, but only 20 of the 42 studies, slightly less than half, had a criterion which was related to a school or classroom setting. If the main objective is to predict school performance, the GATB cannot be highly recommended for use in Idaho.

The Phi Coefficients in Table II will be evaluated according to the correlation levels previously adopted by the writer. These were a correlation of .45 or more for general use, and .30 as a reject level below which no instrument should be used as the sole basis for guidance, selection, or placement. Ten, or nearly one-fourth, of the 42 phi coefficients fall below the .30 level. Fourteen, or one-third are .45 or higher. These results alone indicate that the GATB should be used in Idaho only with extreme caution, but the contents of the paragraphs immediately following cast even further doubt on those studies reporting phi coefficients of .45 or higher.

Variations in Predictive Studies

The dangers inherent in the generalization of predictive studies based on small samples and from one geographic area to another already have received attention. The topic is sufficiently important to merit additional consideration. Table II yields data for another important comparison, namely, the phi coefficients for two or more studies relating to the same specific occupation. There are ten of these, and for each specific occupation, the highest phi obtained, and the lowest phi, are as follows:

Specific Occupations	Highest phi	Lowest phi
Office-Machine Serviceman	.50	.40
Programmer, Business	.35	.31
Dental Assistant	.66	.22
Draftsman (various types)	.30	.23
Electronics Technician	.52	.36
Instrument Repairman I	.35	.26
Stenographer, Typist	.47	.22
Surgical Technician	.53	.28
Patrolman	.30	.22
Nurse, Licensed Practical	.57	.27

A few quick glances up and down the columns of phi coefficients reveals a shocking discrepancy in all but a few cases. Based on the highest phi coefficient, the majority of these validity studies would seem useful in a new predictive situation. One could not be sure, however, that their predictive value would not be more like the results in those studies reporting the lowest phi. This much variation should not be tolerated, and local studies should be completed to ascertain just what the predictive relationship is before the GATB or any instrument is adopted for general use.

Predictive Studies for Training Groups

A good example of what happens when previously developed specific norms on the GATB are applied in a new situation is a recent article by Droege (1968). Data for 875 trainees in programs under the Manpower Development and Training Act (MDTA) were available for 12 different samples. Although the GATB norms showed significant cross validity in 10 of the 12 samples, it is of interest to compare their predictive value with that found in other studies.

Table III on the next page presents the training groups and the phi coefficients reported in Table 4 in Droege's article. The writer obtained the highest phi coefficient from the technical manual (U.S. Department of Labor, 1967a, Table 9-1) for the same or a similar specific occupation. The latter are listed in the third (right-hand) column of Table III. None of the phi coefficients obtained for the training groups were as high as .45; thus the writer would conclude that the GATB did not show sufficient predictive power in any of the 12 samples to justify its use as a selective device in training (the MDTA were not test-selected in these studies). Actually, 5 of the phi coefficients fell below .30, which is nearly half of the 12 sample groups, and most of the others are not much higher. The differences between the phi coefficients for training and the highest phi obtained in other studies follow the pattern already observed above under the previous heading. Whereas more than half of the highest phi coefficients were above .45 and thus suitable for prediction according to the writer's standards, none of the cross validation studies reached this level.

DISCUSSION, SUMMARY, AND CONCLUSIONS

Discussion

This chapter has not been too kind to the GATB. The writer may be suspected of setting out originally with the sole purpose of discrediting this test battery. The reader is reminded that the writer stated his premises clearly. Moreover, selective criticism was in order, as the basic purpose was to relate the research reported for the GATB specifically to Idaho's system of vocational-technical education. Most important, this was accomplished by taking a careful look at the research data that employees of the U.S. Department of Labor themselves obtained and published. The reader is free to go to the same sources and draw his own conclusions.

Most sections of this chapter do not need further elaboration, as their contents speak for themselves. A few merit further discussion, the first of which is the test-retest variability of the GATB. A possible solution to this problem would be to administer tests at different intervals over a student's school career. Better yet, students should have an opportunity to choose the particular day they

TABLE III

PHI COEFFICIENTS FOR TRAINING GROUPS COMPARED WITH THE HIGHEST PHI COEFFICIENTS OBTAINED IN EARLIER STUDIES WITH THE GATB

Training Group	Training phi	Highest phi
Auto-Body Repairman	.34	.55
Automobile Mechanic	.26	.40
Auto Service Station Mechanic	.34	.45
Clerk-Stenographer	.31	.47*
Clerk-Stenographer	.25	.47*
Diesel Mechanic	.00	.33
Hand Sewer, Shoes	.37	.52
Instrument Repairman	.26	.35
Operating Engineer	.14	.24
Production Machine Operator	.38	.38
Stenographer	.42	.47
Tractor-Trailer-Truck Driver	.32	.50

*Phi is for Stenographer-Typist, as no entry was found for Clerk-Stenographer will be tested, when they feel they can do their best. Retesting might be in order at the request of some students, and even permission not to be tested at all if this is what they prefer. Most administrators and other school officials probably would throw up their hands at such suggestions and say there just isn't time. The writer merely would invoke his earlier premise, and suggest that the testing might better be omitted altogether.

It is evident that local validation studies are badly needed. The variations noted in studies with the GATB are not unique in this respect. Ghiselli (1966, p. 28ff.) noted similar wide variations, and the manual for the Differential Aptitude Tests (Bennett, et al., p. 5-2ff.) also reveals important differences in correlations obtained in separate validation studies. The point to be emphasized once again is that studies completed outside Idaho and with varying samples cannot be applied directly to Idaho students without the likelihood of important changes in predictive power. Local validation studies obviously require professional research time and funds, and it is understandable that so few studies have been completed. This writer maintains they should be provided for, or that the use of the GATB be restricted to use by trained counselors only as a guidance tool.

The latter statement needs further elaboration, as it may seem inconsistent with the earlier discussion. The writer is saying that the GATB may be used by trained counselors as one tool (or means) in vocational guidance, but that the GATB cannot be recommended for use in selection and placement. The reason for this apparent inconsistency should become clearer in view of the writer's conception of the role of a trained counselor. He will exhibit at least the following:

--He will help the counselee to understand that, at best, considerable uncertainty still remains in career planning regardless of which, and how many, tests are administered and interpreted.

--He assumes that the best choice for the counselee is the one for which the counselee is willing to accept full responsibility.

--The trained counselor will be thoroughly familiar with the technical aspects of the appraisal instruments he uses, such as reliability and validity, as well as their relative strengths and limitations.

--He will not slavishly follow a multiple cut-off method, or any other statistical method, of test interpretation. Moreover, he will attempt to understand the psychological continua which underlie the test performances of the counselee.

--Other tests in addition to, or instead of, the GATB will be administered and interpreted when their use is indicated.

--The counselee will be permitted, and perhaps even encouraged, to choose a career objective which isn't necessarily congruent with the best predictions from a test appraisal.

- --The trained counselor will assist the counselee to integrate all facts about himself, and relate these to a career objective which appears to maximize his own chances of success and satisfaction, and his contributions to society.

The counselor, in other words, focuses on the welfare of the individual counselee, and uses the GATB or any test as a means to this end. Selection and placement tend to be more impersonal and thus more dependent upon proven predictive relationships between tests and success criteria in a specific setting. The GATB thus is recommended as one tool in the counselor's kit, but cannot be recommended as a selection and placement tool without additional research in the setting in which it is to be used.

Summary

The General Aptitude Test Battery (GATB) is probably the best known multi-factor battery in the vocational area. It has been used in a vast number of research studies, and likely is unsurpassed in regard to the technical data which are on hand. Despite this wealth of research, the GATB was shown to have a number of deficiencies. Of 425 phi coefficients obtained in separate validation studies, 258, or 61 per cent, fell below .45, a minimum level the writer considered necessary for practical use. Phi coefficients less than .30 numbered 74, or 17 per cent, a range the writer considered unsatisfactory for any practical application. Test-retest variability was noted, and research studies were suggested which might reduce this or take it into account. Moreover, test administration policies were suggested which might better account for test-retest variations. Specific comparisons were made between research data for the GATB and vocational programs in Idaho's area vocational schools. The results indicated that the GATB could be recommended only as one tool in the kit of a trained counselor. It could not be recommended for selection and placement in a local setting without additional research.

Conclusions

Based on the contents of this chapter, the following conclusions may be stated:

- --The General Aptitude Test Battery, even though it has been researched more extensively than perhaps any other test in the vocational area, can be recommended only as one tool in the kit of a trained counselor. It cannot be recommended for selection and placement unless additional research studies are completed in the specific setting in which the test battery is to be used.

- --Research is needed which bears upon the problem of test-retest variability. Differential prediction studies and a search for moderators which predict individual error of measurement and error of prediction are indicated.

- --Policies concerning the administration of appraisal instruments need to be examined carefully, and changed if necessary to insure the utmost accuracy in the results obtained. The practice, all too common, of administering important tests at one point in time perhaps should be discontinued.

CHAPTER III

VALIDATION STUDIES CONDUCTED BY THE STATE OCCUPATIONAL RESEARCH UNIT

Early in the existence of the State Occupational Research Unit (which was organized officially on June 1, 1965), interest began to develop in the use being made of the General Aptitude Test Battery (GATB) in Idaho's program of vocational-technical education. Less than a year had elapsed before the writer, as Director of the State Occupational Research Unit, had made arrangements with counselors at two of the state's area vocational schools to conduct validation studies. Two of the five existing area vocational schools were selected which had been established for a relatively long period and which enrolled the largest number of students. The area schools selected were those affiliated with Boise State College, Boise; and Idaho State University, Pocatello. As will be revealed in the next chapter, similar studies also have been conducted, although on a smaller scale, at North Idaho Junior College, Coeur d'Alene; and at the College of Southern Idaho, Twin Falls. Two other studies which are discussed in later chapters of this report utilized data from all of Idaho's area vocational schools.

The validation studies reported in this chapter were conducted by the writer. Data were collected during the summer of 1966 with the assistance of the counselors and clerical personnel at Boise State College and at Idaho State University. The original intention was to collect test scores for any and all instruments which had been administered to large numbers of students. A statewide testing program in Idaho secondary schools was initiated in the fall of 1961, composed of the Differential Aptitude Tests and the Iowa Tests of Educational Development. As already indicated, the GATB had been used extensively in the counseling and admission programs of the area vocational schools. Boise State College also had test results for several students on the Nelson-Denny Reading Test. The criteria of academic success were of two basic types, namely, grade-point averages and graduation versus non-graduation.

As the data were collected, it became evident that test results only for the GATB and the Nelson-Denny Reading Test were available in sufficient numbers to justify statistical analyses. A list of 20 different vocational programs was prepared originally, and data were sought for as many cases as possible within each program. The cases also could be classified as graduates, non-graduates, and present enrollees. Statistical analyses were of two basic types. First, zero-order correlations were computed for samples composed of graduates, non-graduates, and present enrollees combined for those vocational programs in which there were a sufficient number of cases. Correlation studies with the GATB were not completed unless data were available for at least 50 cases. Since fewer cases had results on the Nelson-Denny Reading Test, a minimum of 30 cases was adopted for the correlation studies. Analysis of variance studies were completed for several programs in which the graduates and non-graduates were compared on their GATB scores. The latter studies were sometimes based on very small samples. The intent throughout all of the studies was to explore important relationships and differences, rather than to produce a finished system of test-criterion relationships which would have practical utility. The results obtained in the studies at each institution are presented in the next two sections.

RESULTS FOR BOISE STATE COLLEGE

The Correlation Studies

The correlations obtained between the Nelson-Denny Reading Test and cumulative grade point averages in three different program groups are presented in Table IV on the page following. The small numbers of cases preclude any definitive conclusions, but it is apparent that none of the relationships are very high. A correlation of .36 is significant at the 5 per cent level with 28 degrees of freedom (Garrett, 1953, p. 200). Thus, none of the correlations for students in electronics or drafting and design are significant. Since a correlation of .30 is significant at the .05 level with 40 degrees of freedom, the correlations of .39 for vocabulary and .33 for the total score indicate that these relationships are not due to chance for students enrolled in automobile mechanics. If similar results were obtained with a larger number of cases, the Nelson-Denny Reading Test might be a valuable addition to a multiple prediction battery for this one program. The results obtained in this study do not indicate that it would be too useful as a single predictor.

Table V, page 30, presents the correlations between GATB scores and cumulative grade point averages for four vocational program groups. The reader should note that three of the four groups are the same as the three listed in Table IV. A correlation of .27 is significant at the .05 level with 50 degrees of freedom (Garrett, 1953, p. 200), and must increase to .35 to be significant at the .01 level. Smaller correlations are required to meet these significance levels as the number of cases increases, but since the actual magnitude of the correlations is of more concern, the foregoing correlations for 50 degrees of freedom should suffice as a rough guide. A perusal of Table V reveals little predictive value for academic performance in automobile mechanics and in drafting and design. The highest correlation for automobile mechanics was .30 for general learning ability. The correlation of .38 for finger dexterity in drafting and design is notable, and might be an important contributor in a multiple prediction battery. Multiple correlations were not computed in this study because most of the programs had an insufficient number of cases. Judging from his own past experiences, the writer would expect a multiple correlation to be .05 to .10 higher than the highest zero-order correlation for a given program.

TABLE IV

CORRELATIONS BETWEEN NELSON-DENNY READING SCORES AND CUMULATIVE GRADE POINT AVERAGES (BOISE STATE COLLEGE)

Program Groups		Nelson-Denny Scores Vocabulary	Comprehension	Total
Automobile Mechanics	r	.39	.22	.33
(Total N = 45)	N	45	45	45
Drafting & Design	r	-.06	-.08	-.07
(Total N = 36)	N	36	36	36
Electronics	r	.32	.22	.29
(Total N = 30)	N	30	30	30

TABLE V

CORRELATIONS BETWEEN GATB SCORES AND CUMULATIVE GRADE POINT AVERAGES (BOISE STATE COLLEGE)

Program Groups		GATB Scores								
		G	V	N	S	P	Q	K	F	M
Automobile Mechanics	r	.30	.26	.19	.16	-.11	.08	-.08	.17	-.11
(Total N = 107)	N	107	107	107	107	107	107	107	106	106
Drafting & Design	r	.25	.18	.25	.27	.16	.25	.21	.38	.22
(Total N = 121)	N	121	121	120	121	120	119	118	118	117
Electronics	r	.46	.38	.23	.12	.37	.29	.14	.24	.09
(Total N = 57)	N	57	57	57	57	57	57	57	57	57
Welding	r	.64	-.38	.65	-.30	-.41	.40	.37	.33	.32
(Total N = 56)	N	56	56	56	56	56	55	55	54	54

Legend for GATB Scores:

G = General Learning Ability
V = Verbal Aptitude
N = Numerical Aptitude
S = Spatial Aptitude
P = Form Perception
Q = Clerical Perception
K = Motor Coordination
F = Finger Dexterity
M = Manual Dexterity

The results for electronics and welding are both encouraging, and, for the welding students, somewhat bewildering. The writer had the results for the welding students run through the computer on two separate occasions to make certain they were accurate. The predictions for success in electronics, if they held for a larger number of cases, would have practical utility. The correlation of .46 for general learning ability, combined with the other GATB scores in a multiple prediction study, might yield a multiple correlation of .50 or higher. The correlations of .64 for general learning ability and .65 for numerical aptitude for the welding students reflect substantial relationships, which, if they held up in cross validation studies, would yield useful predictions. The correlations of -.30 for spatial aptitude and -.41 for form perception are a little surprising; a positive relationship logically might be expected.

Here is a dramatic example of what may happen when results from another geographic area are applied to students in a training program that seems logically related. Students in this welding course were administered the specific battery for welder, combination (U.S. Department of Labor, 1967a, p. 81, entry 380). The GATB norms are 85 for spatial aptitude, 85 for finger dexterity, and 80 for manual dexterity. The results obtained in this study suggest that the published norms might not be too effective in predicting success for this group of students.

The Analysis of Variance Studies

As noted previously in this chapter, each of the cases for whom data were collected could also be classified as graduates of a vocational program, currently enrolled, or as a non-graduate. It seemed more appropriate to use the latter term instead of "dropout," which perhaps is more common. Such factors as the military draft, personal finances, and others seemed to count heavily in several cases. More important, in some cases students have found employment and have entered an occupation related to their training before completion of the formal training program. This happened, for example, in auto mechanics at Idaho State University with the result that the program was shortened from 2 years to an 11 month program. It seemed more accurate, therefore, to refer to these students collectively as non-graduates, although there may have been some among them who could not succeed academically.

Since it was impossible to determine the ultimate status of those students currently enrolled, they were excluded from the analysis of variance studies. Although some of the non-graduates may have returned and completed their programs subsequent to this study, it seemed logical to group them together and compare them with the graduates. Simple analysis of variance between the means of the two groups yielded results on the number of cases, the means, and the F ratios for each score of the GATB. The analysis of variance studies were completed for the program groups in automobile mechanics, drafting and design, electronics, and welding. Since four tables were required to present the results, they are placed near the end of this report as Appendix A.

The interested reader should turn to Appendix A and examine the four tables after reading the next two paragraphs which discuss their format and contents.

The tables in Appendix A present in the first (left-hand) column the letter symbols for each score of the GATB. These may be interpreted by referring to the introductory page of Appendix A which presents a legend. The next column moving to the right contains the number of cases; the uppermost number is for graduates, and the lower number in each pair is for the non-graduates. The next column presents the corresponding means. The last (right-hand) column contains the F ratios.

Interpretations must proceed cautiously, but at least three trends seem sufficiently clear to deserve comment. First, most of the pairs of means are very similar, indicating little difference in aptitude between the graduates and non-graduates. In several cases, the non-graduates have the higher mean. Second, as might be expected from the similarity of the means, all except one of the F ratios is not significant. Third, there is some tendency for the GATB scores which correlated highest with cumulative grade point averages (see Table V) to show important differences between the graduates and non-graduates. The highest correlations were: automobile mechanics, .30 for general learning ability; drafting and design, .38 for finger dexterity; electronics, .46 for general learning ability; welding, .65 for numerical aptitude. The tables in Appendix A reveal that the means for general learning ability for automobile mechanics are the same. The means differ, however, for the others:

Program Groups	GATB Score	Mean of Graduates	Mean of Non-Graduates
Drafting & Design	F	105.2	97.3
Electronics	G	114.0	110.2
Welding	N	97.0	85.6

The difference in means on numerical aptitude for the welding students yielded the only significant F ratio--5.40, which is significant at the .05 level. These differences are in the direction one would expect from the correlations, and thus add more meaning to the latter. Both groups in electronics have means above average, and since the actual difference is small, the mean difference would not have high practical utility.

RESULTS FOR IDAHO STATE UNIVERSITY

The Correlation Studies

Test scores were available only for the GATB at Idaho State University, but more vocational program groups and a larger number of cases were represented. The results of the correlation studies are presented in Table VI on the next page. The reader may have noticed that, in Table V (page 30) the number of cases (N row) for each program group were the same or only slightly less than the total N. Idaho State University was more inclined to require that the applicant take only those sections of the GATB that were necessary to yield specific norms for the program entered. A perusal of Table VI reveals that there is more variation in the number of cases, especially for finger dexterity and manual dexterity. Since the latter two scores were obtained from performances on special apparatus, there apparently was some tendency to omit these sections to conserve test administration time.

The correlations in Table VI are, for the most part, too small to provide predictions for practical use. Exceptions are the programs of instrumentation, with a high correlation of .49 for general learning ability; and secretarial training with .43 for numerical aptitude as the highest correlation. Academic success in automobile mechanics and in drafting and design was not predicted too well at either Boise State College or Idaho State University (compare Tables V and VI). Electronics success at Boise State College had a correlation of .46 with general learning ability, based on 57 cases. The comparable correlation at Idaho State University was only .26, based on a much larger sample of 241 cases. Both correlations were the highest obtained for the two programs. The differences may be genuine, as the program at Boise State College was a 2 year program in industrial electronics, whereas the program at Idaho State University was more technical, and required 3 years. It also is possible that the lower correlation, based on the larger sample, is closer to the "true" relationship.

The correlations for the last two programs entered in Table VI are a little more encouraging. The highest correlations for instrumentation are .49 and .45 for general learning ability and verbal aptitude, respectively. They are, however, based on slightly over 50 cases, and might change substantially on cross-validation. Similar to the results for welding at Boise State College (see Table V and accompanying discussion), there is an important difference between the high predictors obtained in the writer's study and the GATB norms. The technical manual (U.S. Department of Labor, 1967a, p. 69, entry 176) gives norms for Instrument Repairman I as 85 for numerical aptitude, 100 for spatial aptitude, and 80 for form perception. Only the correlation of .35 for spatial aptitude in Table VI would lend much support to the norms developed in earlier studies.

On the other hand, the correlations for secretarial training of .39 and .43 for general learning ability and numerical aptitude, respectively, compare more favorably with norms developed in earlier studies (U.S. Department of Labor, 1967a, p. 62, entry 37; and p. 78, entry 331). The secretarial program at Idaho State University contained two main emphases--either stenographic and related or bookkeeping. The technical manual gives as norms for Bookkeeper I a 90 on general learning ability, 95 on verbal aptitude, and 95 on numerical aptitude. The norms for Stenographer, Typist are 95 for general learning

TABLE VI

CORRELATIONS BETWEEN GATB SCORES AND CUMULATIVE GRADE POINT AVERAGES (IDAHO STATE UNIVERSITY)

Program Groups		GATB Scores								
		G	V	N	S	P	Q	K	F	M
Automobile Mechanics	r	.24	.20	.24	.16	.18	.31	.01	.18	.30
(Total N = 73)	N	73	73	72	73	72	72	72	71	69
Business Mach. Repr.	r	.33	.25	.22	.25	-.02	.17	-.22	-.16	-.29
(Total N = 52)	N	52	52	52	52	52	52	51	44	47
Cosmetology	r	.25	.02	.32	.18	.13	.03	.20	.27	.20
(Total N = 71)	N	71	71	68	71	71	68	70	47	47
Drafting & Design	r	.15	.17	.26	.03	.05	.09	.16	-.01	.15
(Total N = 77)	N	76	76	74	77	75	67	72	63	60
Electronics	r	.26	.14	.16	.21	.12	.20	.04	.14	.18
(Total N = 241)	N	241	239	237	241	239	229	228	177	178
Instrumentation	r	.49	.45	.28	.35	.09	.13	.02	-.04	.07
(Total N = 52)	N	52	51	52	52	50	49	46	39	36
Secretarial Training	r	.39	.30	.43	.06	.16	.29	.08	.05	.04
(Total N = 279)	N	277	277	267	278	275	277	277	161	160

Legend for GATB Scores:

G = General Learning Ability
V = Verbal Aptitude
N = Numerical Aptitude
S = Spatial Aptitude
P = Form Perception
Q = Clerical Perception
K = Motor Coordination
F = Finger Dexterity
M = Manual Dexterity

ability, and 100 each for form perception, clerical perception, and motor coordination. It appears that the specific norms for Bookkeeper I would be the better selector for students in this program. Additional studies are needed, of course, in which the students are grouped according to the more specialized phases of the program in which they are enrolled.

The Analysis of Variance Studies

The tables containing the results of analysis of variance studies for vocational programs at Idaho State University are presented in Appendix B of this report. Similar to the studies reported in the previous section for Boise State College, these analyses compared the graduates and non-graduates in a number of programs on each separate score of the GATB. The reader should examine the tables in Appendix B, as this will improve his understanding and appreciation of the material which follows. The first page of Appendix B presents a legend for the GATB symbols which are listed in the first (left-hand) column of each table. The number of cases form the next column to the right, with the upper number representing graduates and the lower number representing the non-graduates. The means of each group follow, and the F ratios form the last (right-hand) column.

The vast majority of the means are very similar, and in several instances it is the mean for the non-graduates that is the higher of the two. Accordingly, the vast majority of the F ratios are not significant. Several significant F ratios were found, however, and it is instructive to compare the means of the two groups:

Program Groups	GATB Score	Mean of Graduates	Mean of Non-Graduates
Cosmetology	F	115.5	102.3
Electronics	G	114.8	110.3
	S	122.4	114.8
Instrumentation	V	102.9	84.8
Secretarial Training	Q	116.8	112.6

The means for cosmetology and instrumentation are based on 15 and 5 non-graduates, respectively, and thus are only suggestive of important differences which might be found. Although the correlation (see Table VI) for finger dexterity in the cosmetology program is low (.27), the correlation for verbal aptitude in the instrumentation program (.45) is fairly substantial. The mean differences for the remaining two programs compare with correlations in Table VI which were all below the .30 level. Perhaps analysis of variance studies should be completed in addition to correlations in test validation studies; the former may yield important insights which are not revealed when only correlations are available. It is obvious in this portion of the study, however, that the mean differences for electronics and secretarial training are not large and therefore of little practical significance.

DISCUSSION, SUMMARY, AND CONCLUSIONS

Discussion

Several important considerations need to be stated and discussed. The first pertains to the general level of the correlations reported in Tables IV through VI. A few quick glances at each table may leave the impression that most of the correlations are so low as to be worthless for practical prediction. As was pointed out in the previous chapter, GATB norms for specific occupations are based on a minimum of two and a maximum of four of the nine aptitude scores. It is not too likely, in view of the different aptitudes measured by the nine GATB scores, that all of them would correlate highly with success in any given program. This seems evident from a summary of occupational validity for the GATB, based on 424 studies involving over 25,000 employees, applicants, trainees, and students (U.S. Department of Labor, 1967b). The grand median validity for all aptitudes was a .22 correlation. The picture is brightened considerably, however, when the predictive values of specific batteries of two to four scores are examined. These correlations have a median value for different broad classes of criteria from .40 to .45, which is near the minimum adopted by the writer for practical prediction (Cf. Figure 1, p. 13 of this report).

It also is important to remember that the studies reported in this chapter were based on cases which, to some extent at least, had been test-selected. The writer learned that not everyone admitted to either Boise State College or Idaho State University was required to pass the specific norms on the GATB corresponding to the program in which he or she enrolled. Some selection did occur, however, and the very fact that the GATB was required may have discouraged some potential enrollees. These conditions might tend to lower some of the correlations obtained. But by the same token, the few correlations which were significant, and those reflecting relationships high enough for practical use, can be given serious consideration.

A few correlations for single scores of the GATB were high enough to meet the writer's standard of .45 for practical prediction. These correlations were .46 for general learning ability in electronics, and .64 and .65, respectively, for general learning ability and numerical aptitude in welding at Boise State College; and .49 and .45, respectively, for general learning ability and verbal aptitude in instrumentation at Idaho State University. Assuming that a multiple correlation would add .10 to the highest zero-order correlation, the following additional programs might be added: drafting and design at Boise State College; and secretarial training at Idaho State University. This is 5 programs out of 11, or less than half, for which the GATB could be expected to make practically useful predictions. More important, it was demonstrated that the norms developed in earlier studies with the GATB for welding and instrumentation did not coincide with the scores in the present study which yielded the highest zero-order correlations. The writer expects that test development officials in the U.S. Department of Labor will object to these comparisons, and point out that the writer did not follow their procedure for development of norms for specific occupations (U.S. Department of Labor, 1967a, pp. 43-56). The writer would readily agree, and merely point out that not everyone agrees that the procedures followed by the Labor Department officials are the best available.

The analysis of variance studies added an extra dimension to the investigation of important relationships between test performances and academic success. Two notable findings were that (1) the differences tended to support the results obtained in the correlation studies, but (2) that the two groups generally were quite comparable. Studies are needed to determine why students with aptitudes sufficiently high to indicate success in vocational programs do not remain until graduation. Since it apparently is not primarily a matter of low aptitude, we may need to use other measures of interests, attitudes, personality, etc. to understand this behavior (Cf. Chapter I, pp. 5-7).

The picture that emerges is that the GATB has some potential as a practically useful predictor in Idaho's system of vocational-technical education. Examples have been presented once again which show the necessity of local validation studies. The GATB is relatively simple to administer, and its wide range of measurements makes it worthy of consideration as part of an experimental battery. It is not, of course, the only test available, and certain predictions possibly could be improved by the addition of the Nelson-Denny Reading Test, or others. Some persons in Idaho suspect that the GATB is not a good predictor for high level programs like electronics and drafting and design because these students may tend to "score out the top." That is, the GATB does not measure as well at the higher ability levels as it does at levels closer to the average. Another consideration is the prerequisites for certain Idaho programs. A student is required to have one year each of algebra and geometry before he enters drafting and design, electronics, industrial electronics, and instrumentation. An achievement test might be a more useful predictor for these programs. Since the next chapter presents several studies, some of which compare GATB predictions with those of other tests, further discussion may be found in the last section of Chapter IV.

Summary

Correlation studies and analysis of variance studies conducted by the writer at two area vocational schools in Idaho were presented and discussed. The tests used were the Nelson-Denny Reading Test and the General Aptitude Test Battery. The criteria were grade point averages and graduation versus non-graduation. Both tests yielded significant correlations with grade point averages in 1 or more of the 11 programs studied. Only the GATB correlated high enough to indicate practical application, but this occurred in only 5 of the 11 programs. The analysis of variance studies tended to support the relationships reflected in the correlation studies, but generally the graduates and non-graduates were quite similar in their performances on the GATB.

The need for local validation studies once again was demonstrated and emphasized.

Conclusions

Based upon the contents of this chapter, the following conclusions may be stated:

--The General Aptitude Test Battery, in view of its wide range of measurement and validation research, should be given high priority in an experimental battery. It should not, however, be the only test used or considered in studies of academic prediction and group membership.

--Norms for specific occupations from the General Aptitude Test Battery, even though logically related to a local vocational program, do not always compare favorably with test-criterion relationships based upon student performance in the local program.

--The General Aptitude Test Battery has some potential as a differential predictor between graduates and non-graduates in vocational programs, but its potential in this area needs to be compared with that of interest, attitude, and other similar measures.

CHAPTER IV

A REVIEW OF RESEARCH IN IDAHO

An attempt was made to locate and review all research studies in Idaho which included the use, or development and use, of some instrument for either understanding or predicting vocational behaviors. Research completed during the last 10 years was the primary focus, although a few important studies were included which were completed more than 10 years ago. The writer was assisted by several persons throughout the state to identify some 30 completed research studies. This became more involved than was expected, as several studies concerned with problems in Idaho were completed as graduate theses at institutions in other states. A "reasonable effort" thus was made to identify all pertinent research; it is possible that a few studies were missed. Moreover, it was intended originally that both completed and ongoing research studies would be included. Due to difficulties encountered in identifying ongoing studies, and for other reasons, this review is limited to completed research. It also is limited to those projects which resulted in a written report which contains an account of the procedures followed and the results obtained.

Several studies which dealt with the guidance program in Idaho schools and with the role of the counselor were reviewed, but these were not included in this chapter. The focus in this report is on instruments of one kind or another, and not on the more extensive role and function of the counselor and the total program of guidance. An obvious type of study sought for this chapter is one which determines student characteristics by means of one or more psychological tests and presents certain statistical analyses which aid in understanding and predicting vocational success. The writer also included studies in which a questionnaire was developed as a follow-up instrument, or as a means to obtain information regarding such behaviors as the choice of a college major field of study. Studies of this type were sought because Ghiselli (1955, p. 111ff.) reported several fairly substantial relationships between personal data and vocational success criteria. A personal data blank or a questionnaire thus has potential as an instrument for studying vocational behavior.

Two important considerations need to be discussed before the reader begins a somewhat lengthy excursion through a sizeable number and variety of research resumes. First, it is difficult to do justice to the full content of a research report in only a few paragraphs. A brief prepared by another writer likely would highlight different aspects of the study which to that writer seemed most important. References are given, and the reader is free to examine the completed reports if he is so inclined. Indeed, this is highly recommended if research is being planned which bears on some of the same problem areas. Second, and most important, the reader of this report needs to understand the underlying motive of the writer in criticizing a study. Probably no one who has completed a thesis project or other research comparable in scope emerges from the experience without considering the finished product a "piece of himself." This also may be true of the major professor who has guided a student through his thesis research. The writer is interested in promoting additional research, and one way to accomplish this is to point out deficiencies in completed projects, or problem areas which need additional study. Many of the authors are personal acquaintances of the writer. The writer even included his own thesis research with appropriate criticisms, although he may be suspected of going a little easier on himself than someone less emotionally involved. The studies, generally speaking, are commendable pieces of work and this is taken for granted. Critical remarks thus may appear to outweigh those of a more favorable vein. The hope and intent is that this will stimulate more and better research in the future.

The section which follows contains a brief resume of each research study. Since each study may be identified with one or more persons, the name(s) of the author(s) will appear in the side heading. A date is given in parentheses which is a reference to the corresponding entry in the bibliography. The studies were arranged alphabetically by author and this is the order of presentation. The amount of space devoted to each study is not intended to have any logical significance. Some studies, due to a variety of reasons, are easy to summarize in a few words. Others need more space in order to convey an understanding of their design and results.

IDAHO RESEARCH STUDIES

The Ayarra Study (1967)

Ayarra investigated the relationship between self-concept and vocational preference. Three hypotheses, stated in null form, were tested:

1. There will be no significant difference in vocational preference between students with a positive self-concept and students with a negative self-concept.

2. There will be no significant difference in means in vocational preference between boys with a positive self-concept and boys with a negative self-concept.

3. There will be no significant difference in means in vocational preference between girls with a positive self-concept and girls with a negative self-concept.

The operational measure for the self-concept was the self-regard score, Column I, of the Bills Index of Adjustment and Values. The operational measure for vocational preference was the student's response to a questionnaire modeled after Super's three-way classification of occupations. The student was presented a list of occupations, each of which was previously classified according to the three-way model. The range was from professional (level one) through unskilled (level six) occupations.

The basic sample was composed of the senior class (N = 164) of the Preston High School, Preston. Sub-groups were selected from the basic sample on the basis of their scores on the Bills Index of Adjustment and Values. All persons scoring above the third quartile and below the first quartile were selected, composing a high self-concept group and a low self-concept group, respectively, of 35 students each. The low self-concept group contained 16 females and 19 males; the high self-concept group contained 13 females and 22 males. Neither the high and low total groups nor the high and low groups divided according to sex differed significantly on their performance on the National Education Development Tests. The groups did differ very significantly, however, on the self-concept measure. This might be expected, as this was the basis for their selection as extreme groups. Moreover, these results were fortuitous, in that comparisons of significantly different groups on the self-concept measure would increase the likelihood that significant differences on other variables would emerge.

The most important analyses, designed to test the three hypotheses stated at the beginning of the study, consisted of comparing the groups on the measure of vocational preference. The tests were made with two-tailed t-tests for small samples. Differences significant beyond the .05 level were found between high self-concept males and low self-concept males ($t = 2.95$) and between the total high and low groups ($t = 2.02$) with both sexes combined. The difference between the vocational preferences of high self-concept females and low self-concept females ($t = .72$) was not significant. Ayarra thus was able to reject hypotheses 1 and 2 (see above), but failed to reject his third hypothesis.

An obvious characteristic of this study was that the behaviors which were compared were all verbal-type responses to the instruments selected for, or developed for, the study. The crucial question is: how do these differential responses relate to actual behaviors at a subsequent date, such as college entrance versus direct entry into the world of work, types of work entered, career pattern development, and others. Another feature pertains to the use of the term self-concept. Examination of Super's monograph (Super, *et al.*, 1963, pp. 17-32), which Ayarra also cites, reveals that the self-concept is tremendously complex. Indeed, some researchers interpret it as virtually synonymous with the total personality. It seems a little presumptuous, therefore, to derive a measure from an adjective check list and refer to it as the self-concept. Ayarra *does* recognize this distinction, but several statements lead the reader to think of the measure derived from the Bills Index of Adjustment and Values as a self-concept measure. Studies of achievement motivation might be expected to yield a similar result when compared with vocational preferences classified by levels. A label of "need: achievement" for the self-regard score of the Bills Index would have been more modest (and possibly more accurate) than "self-concept."

The Bertrand Study (1966)

One purpose of this study was to assemble those reasons for school dropout given by writers on the subject, and to compare the findings with those recorded for pupils who withdrew from the Blackfoot, Idaho High School during the 1961-62 and 1963-64 school years. Another phase of the study which is of most interest here was the development of an attitude rating scale based on the semantic differential technique. The instrument was administered to 196 senior class members who enrolled in 12th grade English during the 1964-65 term, and a number of subgroups were formed and compared. A total of 15 concepts were measured, based on 5 polar pairs of adjectives for each concept. A reliability check yielded significant correlations at the .01 level between 67 of the 75 adjective scales administered about one semester apart.

Certain findings which seemed most important to this writer will be stated. The results were presented for comparisons between nine paired sub-groups, and constituted an almost bewildering array of specific findings. A total of 61 students were classed as dropouts. The dropouts tended to be past the legal age limit, thus refuting the common assumption that they were just waiting until they were old enough to quit school. The reasons given for dropping out were highly varied. Another notable finding was that the male dropouts did not indicate attitudes of antagonism toward authority as embodied in the school, home, church, or the law. More study was suggested to get at the "real" reasons for student dropouts.

This study reflected a considerable amount of thought and effort on the part of the experimenter. It seems logical to expect that student attitudes play some part in their decision to leave school, and the semantic differential appears to be one useful method of measuring such attitudes. Several questions arise, however, regarding the study design and the results obtained. The reliabilities, although mostly significant, were rather low. The results were not always easy to interpret, and may have reflected temporary idiosyncrasies rather than lasting attitudes. If an ultimate objective is to predict school dropouts, the results obtained from the measurement of attitudes would need to be compared with other predictors such as poor reading skills, grade failures, etc. Repetition of this study is badly needed before any of the findings can be given serious consideration. Perhaps the greatest need (at least in the view of this writer) is for the entire study design to be related to some theory or an experimental design which clearly reflects an overall purpose or strategy. As it is, many of the interpretations were ad hoc and thus difficult to integrate into a meaningful whole.

The Brenna Study (1961)

Student occupational interests and aspirations were the subject of this study by Brenna. Senior students in 25 selected high schools were administered a specially designed questionnaire, and their responses were compared with potential openings in the Idaho labor force. Results from the student questionnaires were compared with the results obtained by the Employment Security Agency (now Department of Employment) in occupational surveys in the same communities. Returns in both phases of the study were excellent, with about 90 per cent of the questionnaires being received both from the students and from employers. The numbers of cases were sufficiently large to add meaning and stability to the results. A total of 2,935 student questionnaires (1,418 males and 1,517 females) were received. The employing establishments responding to the questionnaire had a total employment of 106,768--and 30,320 of these were women.

Brenna found several discrepancies between the occupational selections of the students and occupational opportunities in their home communities. For example, numerous students expressed a desire to enter the professions, whereas opportunities available in the labor force indicated that relatively few would be able to realize their ambitions. Skilled trades and technical fields offered work opportunities, but relatively few students revealed an interest in entering these occupations. The findings indicated a need for guidance services. Apparently the existing guidance services were exerting only a minor influence on the expressed occupational choices of students.

The results of this study are not unlike several others of this type which have been conducted. It is quite common to find large numbers (perhaps more than half) of a group of high school students who express a desire and/or intention to enter the professional field. Some studies have dealt with this tendency by asking two distinct questions. First, the students are asked to express their ideal occupation. Second, they are requested to indicate the occupation they really expect they will enter upon completion of their schooling. The latter question tends to produce a more realistic picture. Brenna did not make this distinction in the construction of his questionnaire.

Another area which deserves comment is the method used to compare the occupational choices of the students with the results obtained in the surveys conducted by the Employment Security Agency. The latter agency used the Dictionary of Occupational Titles to identify the occupations existing in the communities. It is not clear how Brenna coded the responses of the students--whether this was according to the DOT code structure, or in much broader categories. His Table V on page 35 of the thesis suggests that the student responses were grouped into some 10 very broad categories, such as professional, managerial, and clerical. Although comparisons of such broad groupings have some meaning, it seems apparent that much detail is lost. For example, within the professional class several students may have indicated choices which have little representation in the local labor force. Such choices may or may not be realistic; they may reflect out-migration of several students to areas which offer such employment opportunities. Studies of this type would have little meaning unless follow-up studies could be completed to determine how congruent student's expressed choices are with actual subsequent employment and how many enter employment outside their own community and state.

The Brown Study (1968)

A questionnaire was developed with the help of four counselors and two other teachers, and was administered to a class of 136 seniors in the Mountain Home High School. Usable questionnaires were obtained from 121, or 89 per cent, of the students. The questionnaire was fairly comprehensive and covered such areas as educational plans following graduation, vocational plans, and evaluations of the preparation received in high school. The survey was the first part of a two-part study. The second part will consist of a follow-up study during 1969 of the students who graduate. The findings were presented as frequencies and percentages of response to various questions. Examples of the findings were that more than half (52.4 per cent) of the seniors planned to attend a four-year college, whereas 14 per cent planned to attend a technical or trade school. Most of the students were satisfied with their high school educational program, but several expressed a desire for more elective courses, particularly in fine arts and vocational training.

This study is superior to most high school follow-ups this writer has encountered in Idaho. The collection of data from students both before and after they graduate should enable a researcher to investigate several important relationships. The questionnaire developed, however, may not be the best that could be devised for the study. It contains several questions which are relatively "open-ended," thus requiring the researcher to make several analyses and interpretations of responses. The latter can be handled adequately if some system for interpretation is worked out in advance, but Brown seems to group several responses in an *ad hoc* manner. This does not facilitate interpretation from one year to the next, or from one school to another. Perhaps more structure should be built into the questionnaire, and this also would make possible more rapid tallies of the responses. A final observation is that no attention is given to whether the differences observed are significant statistically. The significance of the difference between percentages, or even an arbitrarily chosen level for interpreting certain differences, would add much more meaning to the results obtained.

The Burke Study (1967)

Burke investigated the predictive efficiency of the General Aptitude Test Battery (GATB) relative to success in a vocational program of industrial electronics. Fifty male students in either the first or second year of the industrial electronics program at North Idaho Junior College, Coeur d'Alene composed the sample. All had taken the GATB or portions of it as part of the entrance requirements. The criterion (dependent variable) was a cumulative grade point average. Twelve of the subjects had finished one year of training, and 38 had completed the two years required for graduation. Statistical analyses yielded means, standard deviations, product-moment correlations, and multiple correlations.

The results for the means and standard deviations revealed that, generally speaking, the sample tended to exceed the national norm group for the GATB and to exhibit less variability. In other words, the group tended to have above-average aptitude and to be more homogeneous than the norm group. The following correlations were obtained between factor scores of the GATB and cumulative grade point averages:

GATB Factor Score	Pearson r
General Learning Ability	.348*
Verbal Aptitude	.158
Numerical Aptitude	.361*
Spatial Aptitude	.111
Form Perception	-.033
Clerical Perception	.272
Motor Coordination	.104
Finger Dexterity	-.045
Manual Dexterity	-.010

The two starred correlations are the only ones significant at the .05 level. A multiple correlation of .386, significant at the .05 level, was obtained by combining general learning ability and numerical aptitude. An F test revealed that this increase was not significant when compared with the Pearson correlations for the single factor scores. A combination of general learning ability, numerical aptitude, form perception, verbal aptitude, and spatial aptitude yielded a multiple correlation of .411, a negligible increase.

This is a good study, although it is restricted in range. The design is sound, in that data for the independent variables were collected prior to training and were related to an index of subsequent academic performance. Similar studies with larger samples are needed for the more than 30 other vocational programs currently offered in Idaho. Burke recognizes and discusses the possibility that the low variability of the sample on the aptitude factors, plus their tendency to be above the average of the national norm sample, may have attenuated the correlations. Even if no selection occurs based on test scores, it may be expected that the mere requirement of a test battery, and the reputation of electronics as a fairly difficult course of study, will tend to discourage the enrollment of low-aptitude students. If selection is carried out based on GATB scores, studies which report correlations between aptitude factors and academic success criteria are likely to underestimate seriously the predictive potential of this test battery.

This study investigated the predictive efficiency of only one test battery. Future studies should incorporate other measures such as achievement tests and biographical information. The criterion side of the design also deserves attention. Burke used a criterion in the form of school grades which has been almost too common in studies of this type. Other less obvious academic criteria, such as special honors and differential patterns of course achievement, might be included in future studies. One also wonders how both aptitude test scores and grade point averages relate to more ultimate criteria of occupational success once the students have departed from the school setting. The difficulty of designing such studies and collecting the necessary data should not close our eyes to their importance for vocational guidance and counseling.

The Cain Study (1969)

The vocational choice process has been the focus of both theory construction and a number of empirical studies. Cain investigated curricular and occupational choice as it related to the profession of mining engineering. Certain propositions were selected from Super's theory of vocational choice, and hypotheses were derived for testing. The data were collected by means of commercially available psychological tests and a locally designed questionnaire. Data were collected in five major areas: (1) interests were measured by the Kuder Occupational Interest Survey, Form DD, (2) the Gough Adjective Check List provided operational measures of the self concept, (3) a questionnaire section presented 22 occupational titles which were ranked according to personal preference, (4) a questionnaire section contained 10 job factors which were rated in a paired comparisons format, and (5) a questionnaire section which was composed of an information test of 37 true-false statements concerned with the occupation of mining engineer. A combination of purposive and random sampling was employed to select an all-male sample of 1,375 high school sophomores, 380 high school seniors, and 58 college undergraduates enrolled in mining engineering.

The results supported Super's theory in several instances, but not so clearly in others. Interests were operative in the choice process, but were not the only determinants having a significant influence. Perceptions of self differed from perceptions of the typical person in mining engineering. This was particularly true for the high school seniors. The occupation of mining engineer was ranked lowest among engineering occupations by the high school students, whereas the students in mining engineering tended to give it a high rating. The job factors concerned with social aspects and intrinsic aspects of a job tended to receive the highest ratings. These factors might rival or exceed interests in importance in the vocational choice of several persons. As might be expected, the mining engineering students scored significantly better on the information test. The high school seniors and sophomores did not differ significantly.

This study excells all others completed in Idaho which were concerned with choice and recruitment processes relative to a field of work or study. The design, data-gathering instruments, and statistical analyses reflect considerable research expertise. Three areas may be identified, however, which might be considered weaknesses in the study. First, the basis upon which certain hypotheses were rejected was not made entirely explicit. Significance levels were stated explicitly, but certain group comparisons were based, for example, on the 24 scales derived from the Gough Adjective Check List. A certain number of significant differences might be expected to occur by chance in several of the comparisons made. A relatively small number of significant differences thus may not be a sufficient basis for rejecting a null hypothesis. Second, the application of certain theoretical propositions to explicit findings in the study seemed ambiguous. This probably stems from the very general nature of most of the theoretical propositions. Even so, the interpretations made were not always free from alternative interpretations. Finally, more interpretation and discussion of certain findings regarding the psychological and sociological processes which were operative would have been desirable. For example, the 24 scales derived from the Gough Adjective Check List could have been interpreted individually and collectively to yield insights into why high school students tend to reject the occupation of mining engineer.

The Chatburn Study (1956)

This was a questionnaire follow-up of graduates of Boise Junior College (now Boise State College) from 1934 through 1954. The number of graduates during this 21-year period was 1,576, of whom 43 per cent were women and 57 per cent men. Based on the total living graduates (1,543), the 1,009 returns represent a 65.4 per cent response. Addresses could be found for 1,395, and of this number 72.3 per cent responded. The questionnaire was very extensive, and a vast amount of information was received and analyzed. It appears that the experimenter was mainly interested in the academic side of the curriculum, as relatively few questions pertained to the vocational programs offered at the College. Instruction in secretarial science, cabinet making, sheet metal, automobile body and fender, automobile mechanics, and machine shop were offered at this time.

One part of the survey was concerned with ascertaining the extent to which the college was meeting the needs of those students and other graduates who did not transfer to a senior college. Accordingly, the question was asked: "If you began work immediately after leaving Boise Junior College, do you feel that the school had given you adequate preparation for your job?" Graduates answering in the negative were invited to indicate in what ways they felt their training had been inadequate. A total of 366 alumni responded to this question, and 327, or 89 per cent, indicated they felt that they had received adequate preparation. Of the 39, or 11 per cent, answering in the negative, 15 reported that they were following a different kind of work from that for which they had been trained. More than two-thirds of the respondents reported that, if they were starting college over again, they would want courses combining both general education and specific vocational training. Less than one-fourth would take liberal arts courses exclusively, and only about one-eighth would pursue strictly vocational training.

As already indicated, the study seemed to be oriented toward the academic side of the program at Boise Junior College. Accordingly, the results had only general implications for the vocational divisions and were not specific enough to make possible meaningful evaluations of the vocational training received by the students. The questionnaire, while very extensive in coverage, does not appear too suitable as a follow-up instrument for strictly vocational students. Moreover, a few more sophisticated statistical analyses might have been used to test certain relationships and differences. Even so, this was a study which seldom has been surpassed in the scope of information requested from former students and the period of time covered. Researchers in both the academic and vocational divisions of schools or colleges could get useful ideas from this study on questionnaire construction and methods of arranging and analyzing data.

The Faulks Study (1967)

Idaho depends heavily on agriculture and agriculturally-related industries and occupations for the largest percentage of its economic livelihood. Recruitment of agricultural education majors was identified as a problem in recent years and was the focus of this study. There were two phases. First, information was obtained from vocational agriculture instructors in Idaho. A questionnaire was mailed to 77 instructors, and 74 responded, or 96 per cent. The second phase, which is of most interest in this report, involved the distribution of the same questionnaire to agricultural education majors at the University of Idaho. Questionnaires were sent to 56 students, of whom only 22, or 39 per cent, responded.

Faulks states that this was a "disappointing return." This writer would consider it unacceptable for the type of study undertaken. Thus, none of the findings concerning the students can be given much weight. The questionnaire obviously was concerned with agriculture, and was administered only to persons already engaged directly in the field. This type of study can have value in understanding certain behaviors of the groups studied, and provide additional clues as to how potential agricultural education students decide either for or against the field. It seems, however, that an instrument less obviously slanted toward the agricultural field would be needed before a study could be conducted which would have predictive value and which would differentiate agricultural majors from those in other fields of study.

The Fifield-Jones Study (1964)

This is the best example of a research study of high school dropouts in Idaho that has come to the writer's attention. It includes a large number of cases and is exemplary in regard to the methods and scope of data analyses. A survey-experimental design was employed which utilized the questionnaire method to collect data which was compared for dropouts versus a randomly-selected sample of stayins in the Pocatello school system. The study was conducted in two broad phases. The first phase compared students who already had dropped from school between 1959 and 1963 with a sample of students who had been enrolled during the same period. Unfortunately, only 126, or 25 per cent, of the 501 dropout questionnaires were returned in usable form. Only slightly over half (113 of 200, or 57 per cent) of the stayins responded. Therefore, the findings for this phase of the study were only suggestive of important differences between the two groups.

The second phase of the study concentrated on students who enrolled in the Pocatello schools during the 1963-64 term. Those who dropped out were compared with a random sample of stayins. Extensive personal data, results on psychological tests, academic grades, and other types of data were collected as bases for comparisons. Once again, difficulties were encountered in contacting dropouts to the extent necessary for data collections. It was estimated that data were collected on approximately half of the students who actually dropped during this period. An appendix in the report indicates that the analyses were based on 101 dropouts and 143 stayins. Several refined statistical analyses were performed, including chi-square, critical ratio, and correlation.

Only a few examples of the numerous findings reported can be included here. The results revealed many characteristics of the dropouts, a number of which resemble those found in other studies, such as relatively poor performance on achievement measures, less participation in school activities, more grades repeated, and poor reading ability. Several differences between the sexes were revealed. Curricular implications of the findings were stated, including the need for career information and guidance, and for additional course offerings in terminal or vocational types of training. Many of the practical problems encountered in research in the school setting are described and discussed in the report. It should be an excellent source for those planning similar studies.

The Fifield-Watson Study (1967; Watson, 1967)

Graduates of the high schools in Idaho Falls and Pocatello were the subjects of a follow-up study which included classes from 1954 through 1963, a span of 10 years. Fifield and Watson were able to obtain current addresses for about 70 per cent of an estimated 8,500 students. Nearly 6,000 questionnaires were mailed, and between 62 and 63 per cent of these were returned. The questionnaire was very comprehensive, and requested information in the following broad areas: personal information, educational experiences, evaluation of high school experiences, personal adjustment, leisure time activities, and occupational experiences. A personal data sheet also was prepared for each student from school records.

Similar to the study by Fifield and Jones (see above) the vast amount of detail in this study almost defies summarization. Accordingly, a few salient findings concerned with vocational experiences will be presented. Nearly half of the graduates attended an academic university, but over 50 per cent of these soon dropped out. Inadequacies were reported in occupational information, vocational training, post-high school job placement, and vocational guidance. Job satisfaction was high and was related more to the nature of the work performed than to the amount of income. A majority of the college educated and technically trained graduates left Idaho, whereas those with little or no post-high school training tended to remain in Idaho--many near their home towns.

Intercorrelation tables, based on as many as 28 variables, were constructed. Several of the variables were scores from psychological tests of academic aptitude and achievement. Certain relationships revealed that future occupational and educational levels could be predicted best by cumulative grade point in high school, followed in order by achievement test scores and third by ability test scores. Important differences between males and females were observed. One example was that the relationships between measures of achievement and educational or occupational attainment was higher for male than for female graduates. This was interpreted as a lesser tendency for the female respondents to make full use of their capabilities in this area of activity. The length of time out of school seemed to affect the intercorrelations. Those graduating during the first five years appeared to have developed and established relatively stable employment and educational levels. The more recent graduates apparently were still in a stage of transition.

This study probably is the most comprehensive follow-up of high school graduates ever conducted in Idaho. This is true not only in respect to the large number of cases, but also in terms of the breadth of data obtained. It was limited, as are all questionnaire studies, to a number of subjective reports by the former students, many of whom had been out of school nearly 15 years. The fact that several thousand former students could not be located or did not respond undoubtedly biased the results. The authors acknowledged these limitations, and cautioned readers along these and other lines as they interpreted the results. Other weaknesses might be identified, such as the methods used to determine and assign occupational levels and job satisfaction levels. It is doubtful, however, if most practical studies of this type could be improved in many respects without vast increases in the costs involved.

The Flores Study (1966)

This study investigated, not the ability of students to make occupational choices at the eighth grade level, but their ability to choose a curriculum of study for their high school career. The specific purpose was to investigate level of occupational aspiration (LOA), one facet of the occupational choice-making process. The assumption was that LOA must be formed in eighth graders if they are to be ready to choose a high school curriculum which will contribute toward their adult occupational goals. The following hypothesis was tested: There is no significant difference between eighth-grade and twelfth-grade males in the distribution, degree of relationship to I.Q., or stability of their levels of occupational aspiration. The instruments used were the Occupational Aspiration Scale (OAS), constructed by Haller and Miller; and the Otis Quick-Scoring Test of Intelligence, Beta Test, Form CM.

The subjects were male students selected from the Lewiston schools. Samples were selected as follows: 31 of 202 seniors, 33 of 239 juniors, 29 of 119 eighth graders, and 19 of 113 eighth graders.

There were some further reductions due to lack of data. Analyses using significance statistics and correlation were completed. Comparisons of the means and sigmas of eighth graders and seniors on LOA revealed no significant differences. Comparing performances on LOA and IQ, no significant differences were found between eighth graders and seniors. A test-retest comparison on LOA over a six-month period revealed that the eighth graders actually were more stable. Flores concluded that the LOA of eighth-grade students is probably sufficiently realistic and well-formed to allow eighth-grade males to make valid choices of secondary school courses and curricula.

This is an exemplary study considering that it was done at the masters level. It reflects a real attempt to relate research to career development, a research area which has generated lively interest and a respectable amount of research in recent years. Something needs to be said, however, regarding the instrument used to measure level of occupational aspiration. The OAS was based primarily on 90 occupations used in the study conducted by the National Opinion Research Center (NORC) to investigate occupational prestige. There are other studies of occupational aspiration which are based on extensive questionnaires, and these instruments might yield different results than the OAS. The samples Flores used were small, and the findings should be cross validated on larger groups. The findings seem reasonably congruent with those obtained in the Career Pattern Study (Super, et al., 1960) which revealed that ninth-grade boys were ready for exploratory types of experiences relative to the world of work, but were not yet ready to make specific choices.

The Gardner Study (1957)

Lagging enrollments in colleges of agriculture was the incentive for this study which was designed to identify factors influencing farm youth in selecting college curricula. The results were intended to provide some basis for recruitment of agricultural majors and to improve understanding of curricular choice either for or against agriculture. Two separate questionnaires were developed and sent to an all-male group of non-agricultural majors and a group of agricultural majors. The two sample groups were farm youth, so identified from registration cards at the University of Idaho according to the occupation of the father or guardian. The non-agricultural sample had 431 cases, of whom 69.1 per cent responded; and the agricultural sample contained 270 cases, of whom 76.6 per cent responded.

The individual items of each questionnaire were analyzed according to the number of respondents who marked each alternative. Statements were considered a major influence if marked by 40 per cent or more of the non-agricultural majors, and by 80 per cent or more of the agricultural majors. The result was an almost bewildering array of findings. There were some logical interpretations, but several findings were presented with little or no effort to interpret them. Typical of the major influences for non-agricultural majors was that they considered a farmer as tied down to the farm all of the time because of chores that must be cared for every day. The agricultural majors reported that one reason for their choice was a liking for the working conditions associated with agriculture.

This study resembles several others included in this review. A relatively large number of questionnaire items were analyzed to determine characteristics of the groups studied. Simple percentages were used, rather than more sophisticated significance tests. The content of both questionnaires seemed to be slanted toward the group for which it was intended. This might have advantages in a limited study, but the instruments probably would not be suitable for use with a large number and variety of occupational groups. Certain of the findings, based on frequency of response, were logical, whereas others were difficult to interpret. There is a definite need in a study of this type for some rationale or theoretical statement to aid in the interpretation of the results.

The Gentry Study (1967)

Former students, both graduates and non-graduates in three separate vocational programs, were the subjects of a questionnaire follow-up study at North Idaho Junior College, Coeur d'Alene. The program names, the number of questionnaires mailed, and the number and percentage returned, in that order, are as follows: automobile body and fender, 25 and 15, or 60 per cent; automobile mechanics, 70 and 32, or 46 per cent; and industrial electronics, 73 and 54, or 74 per cent. The first two pages of the questionnaire were identical, whereas the third page was written specifically to correspond to the content of the program in which the student had been enrolled. Those graduating during the five-year period 1960 through 1965 were asked to (1) evaluate their experiences at the College and (2) to supply information on their activities since their departure from school.

Results for each program were analyzed separately, but several responses also were fairly common to the three program groups. It was interesting to note that those who responded had more years of previous schooling, more semesters at the College, and higher grade point averages than the non-respondents. Two of the three comparisons showed significant differences for the auto body and fender group. This supports findings in other studies which have revealed that respondents in questionnaire studies may be a biased group compared with the total population of respondents. The respondents generally were satisfied with the training they had received, but indicated that more assistance was needed in job placement. An unexpected finding was that the majority of jobs held were *not* related to the training received. Other results had implications for curricular and program changes.

The small numbers of cases necessitated cautious interpretations, and this was acknowledged in the report. A major purpose of the study, however, was to develop a model for follow-up studies elsewhere in Idaho. Instructional personnel at North Idaho Junior College were involved in the development of the questionnaire and study design, a procedure which would seem most desirable in all similar studies. It is by no means certain, however, that the questionnaire is nearly as refined as it might become with additional development. As noted for the Brown study (see above, pages 43-44), certain data were requested in "open-ended" form, and thus required some interpretation. Different designs and additional data could be incorporated into other follow-up studies which likely would provide more useful and effective information for educational planning.

The Holmquist Study (1967)

A group of students in vocational agriculture were the subjects of this study. Recipients of the State Farmer Degree, conferred on only 2 per cent of a state's membership in the Future Farmers of America (FFA), were contacted by means of a mail questionnaire. Current addresses were obtained for 172 State Farmers who earned the degree from 1955 through 1960. A total of 136, or 81.4 per cent, returned usable questionnaires (the writer calculated this percentage, and found it to be 79.1 per cent). Since they were chosen on the basis of an outstanding supervised farming program and exceptional leadership qualities, the sample obviously was highly selected.

The results were presented in numerous tables; the most complex analysis was the computation of a percentage. The questionnaire was fairly extensive and seemed to reflect good coverage of pertinent areas of information. The respondents chose farming as a career at 48 per cent, and an additional 18 per cent were in other agricultural occupations. About half of those not in farming *per se* reported they would like to enter farming, but a lack of capital or other reasons prevented this. The number of Idaho boys established in farming was comparable to certain other states, but the Idaho boys were higher in the number obtaining education beyond high school.

Studies of highly selected groups are quite useful, and have yielded much important information. Studies in creativity are one example. The latter studies use a wide array of instruments, including several psychological tests. Holmquist's study, based as it was on a mail questionnaire, was limited in the scope of data which could be collected. Moreover, characteristics of a highly selected group are established and take on more meaning if the group is compared with other more "ordinary" groups. Only in this way can certain characteristics peculiar to the select group definitely be established; high (or low) frequencies of occurence may be similar across groups and thus of no value as indices of group membership.

The Johnston Study (1967)

A follow-up study of dropouts in the Pocatello schools was combined with a survey of employers regarding potential job openings and other related occupational information. Questionnaires were mailed to a random sample of two hundred students who left school before graduation from the fall of 1961 to the spring of 1966. Only 63 responded for a return of 31.5 per cent. A list of 859 employers in the area including Pocatello was compiled from the telephone directory. Questionnaires were mailed to each, and 521, or 60.7 per cent, were answered and returned.

The results were so extensive that they are difficult to interpret. The results were tallied, as in many other similar studies, for various items and separately for students and employers. Typical conclusions were that many students were not receiving adequate vocational education, that potential dropouts should be identified at an early age, and that student placement and follow-up were needed. Failure to develop vocational skills appeared to result in limitations regarding job placement and later promotional opportunities.

This study had a number of worthy objectives, but its design and execution left something to be desired. The thesis was more than three hundred pages long, which is a little bulky for a masters thesis. The first 126 pages or so were given over to a survey of literature. The low percentage of returns for the dropouts renders any conclusions very tentative. The data were not presented too well. For example, several percentages were difficult to interpret, and some appeared to add to more than one hundred. Several useful ideas for similar studies could be obtained by anyone who is willing to cover several pages of material. A trained researcher probably would wish to alter both the format and content of the questionnaire and use more refined procedures to analyze and interpret the data.

The Kiler Study (1967)

Kiler developed and tested an instrument designed to determine individual needs. Need lists constructed by Maslow, Murray, Vernon, and Cleeton were examined. A synthesis and fusion of their lists, plus the addition of other needs, resulted in a list of 33 needs for the study. Rather than using the definitions of needs stated by the aforementioned authors, Kiler chose to derive his definitions from Webster's Seventh New Collegiate Dictionary.

The sample had two major origins. One was termed the "general working population" and was composed of approximately 250 persons (about one-third females), most of whom contacted the local office of the Idaho Department of Employment, Boise. In addition, two small groups of graduate students and one of employment counselors were incorporated into the total sample. The second group was a sample of 50 male inmates at the Idaho State Penitentiary, Boise. All persons who contributed data did so voluntarily.

Three instruments were developed and used in the study. The instrument which was used for the main analyses was a Q-Sort deck of 33 cards, each containing the name of one of the 33 needs developed for the study. The examinees did not actually arrange the card deck in Q-Sort fashion as the term is usually understood. Rather, a ranking method was used. The examinee first separated the deck into two stacks, the first representing active needs at the time and the second representing more secondary or latent needs. Both stacks then were ranked separately with the most important need on top and the very least important on the bottom. When recombined with the "active" stack on top, the needs were assumed to be ranked throughout the deck. A second instrument, used only in a study of reliability, was a paired items scale consisting of one hundred pairs of needs. The 33 needs each were paired a minimum of five times with various other needs, and two of the pairs were repeated to determine consistency of responses. The third instrument was a personal data sheet which yielded information necessary to test certain hypotheses, and to obtain additional data to be used in a follow-up study to be completed subsequent to the thesis.

One major hypothesis was stated in null form: that no satisfactory instrument can be developed to select individual needs. This hypothesis was rejected, based on the results of the study. A number of corollaries, also stated in null form, were derived from the major hypothesis, and the results are summarized for each in the next paragraph. The Q-Sort deck for each examinee was classified into one of three levels. The top 10 needs were considered strong, the bottom 10 weak, and the remaining 13 medium. The resulting frequencies of contrasting groups were compared and tested with the chi-square technique. The .05 level of significance was adopted.

Test-retest reliabilities were obtained over a two-week period for the prison sample. The Q-Sort yielded a median of .78 (Pearson product-moment correlation). The paired items scale yielded a median of .94 (rank-order correlation). Corollary hypotheses, and the results pertaining to each, follow.

1. There will be no significant differences in the needs of males as compared to females. This hypothesis was rejected, in that males expressed greater needs for sleep, sex, exercise, health, and prestige; females had higher needs for friendship, beauty, and service to other people.

2. There will be no significant need differences due to various educational levels. Rejection was possible, as persons with college degrees had higher needs for achievement, beauty, and creativity; those with only a high school education or less expressed greater needs for food and employment.

3. There will be no significant differences in needs between married and unmarried persons. This hypothesis was rejected when it was found that married persons were higher on needs for exercise; unmarried persons sons were higher on needs for sex, excitement, prestige, and influence.

4. There will be no significant difference in the needs of the employed as compared with the unemployed. This hypothesis was rejected, although not very convincingly, as only 2 of the 33 needs were significant at the 5 per cent level. The employed group showed stronger needs for food and liquid (e.g. water).

5. Occupational classification will reveal no significant difference when related to the expressed needs of individuals. Rejection of the null hypothesis was based on only one significant difference. Professional and managerial occupations were classed together and compared with a group of laborers, service workers, and new entrants to the labor market. The lower skilled group indicated a stronger need for friendship.

6. There will be no significant difference in the expressed needs of individuals of various ages. This hypothesis was rejected in that persons 35 years of age and older had greater needs for liquid (e.g. water) and service to other people; those under 35 years expressed greater need for excitement, reputation and prestige.

7. There will be no significant need differences between rural and urban reared individuals. Kiler failed to reject this hypothesis, as no significant need defferences were found between the two groups.

8. There will be no significant difference in the needs of male prison inmates and free men. It was possible to reject this hypothesis. The male prison inmates expressed higher needs for sex and conformity; the free men were higher in needs for friendship and service to other people.

This is an interesting study from the standpoint of design. Moreover, several of the need differences between the groups studied seem logical and congruent with other studies available in the psychological literature. Three important considerations, however, need to be voiced relative to the study design and results. First, several of the groups differed significantly on only a very few of the 33 needs. An obvious extension of this study would be cross-validation on different groups. Many of the differences found could easily have resulted by chance. Only additional samples could establish whether these differences are basic to the composition of the groups studied. Second, the study appears too "basic" in many respects. For example, the use of need definitions derived from a standard dictionary does not reflect high psychological sophistication. Furthermore, the examinees were not given the definitions on the Q-Sort cards. Rather, only the word (or a limited number of words) depicting the need were listed. The exact association of the examinee to such terms is of course unknown. This leads directly to the third, and final, point.

Differences between individuals and among groups in their responses to words and other verbal materials is well known in the field of psychological measurement, and has been for at least 40 years. In this respect, then, Kiler's study has given us no new information. More important, however, is whether these differential responses will give us pertinent information beyond themselves. It is not at all uncommon in psychological measurement to find variables which show significant differences but which do not relate to other criteria of practical interest and importance. Admittedly this is a problem area which Kiler did not attempt to investigate, but it nevertheless is one which will need to be faced before his instrument can be recommended for anything other than experimental use. Studies are needed to determine what occupational roles are potential satisfiers of different needs, and conversely how persons with different need patterns perform in a variety of occupational roles. It remains to be seen via additional research whether Kiler's instrument equals or exceeds in value numerous other available instruments for this kind of study and analysis.

The Lee Study (1958)

The personal interview technique was used to follow up 16 former students of the two-year secretarial science program at Boise Junior College, Boise (presently Boise State College). The sample was limited to those who could be contacted in the immediate vicinity of the College. Both the former students and their employers were interviewed, following an interview schedule previously prepared for each. Only those who graduated since 1955 were included, as it was thought that, after a longer time, experience would be more important than school training. Each girl was questioned about the position she held at the time, which skills she had used on the job, and other similar aspects of her work. Each employer was questioned regarding such factors as personality of the employee and the adequacy of her skills.

The results were presented in narrative form as brief case study reports for each girl. Lee tried to incorporate other data such as grades, test scores, and hobbies and interests, but the case study reports did not present any systematic analyses of these data. Certain conclusions and recommendations were stated as, for example, that the College apparently was doing a fairly good job of training these girls. The small sample, however, makes this a very tentative, and almost presumptuous, conclusion. The data apparently were not analyzed and interpreted according to any logical procedure. Case study reports of this type can add meaning and useful detail to a study, but they do not seem adequate as the sole basis for conclusions and recommendations.

The Loudermilk Study (1964, 1966)

This was the only study located which was conducted in an industrial setting. A total of 570 male applicants composed the sample, all of whom desired to obtain employment in the Clearwater Unit of Potlatch Forests, Inc., Lewiston, Idaho. This is a combined lumber and paper mill, employing an average of slightly more than two thousand workers. The basic purpose of the study was to develop a selection test battery. Accordingly, a battery of 23 predictors was assembled for as many applicants as possible, and a variety of statistical analyses were completed. The predictors were derived from four basic sources: the General Aptitude Test Battery (6 scores), a physical fitness battery developed specifically for the study (6 scores), a personality inventory designed to assess the "dependability-conscientiousness" personality factor (1 score), and 10 items from an employee application record. Four criteria, based on performance in all jobs performed over a six-year period, were developed in the form of work efficiency ratings by supervisors, job tenure, time loss (undesirable absences), and industrial accidents.

The relationships between the 23 predictors and the criteria of job tenure, time loss, and industrial accidents were not high enough to be practically important. Therefore, only certain results obtained with the criterion of work efficiency ratings will be reported. The following correlations were obtained:

Physical Fitness Battery	Pearson r
Two-minute sit-ups	-.01
Dip-ups on parallel bars	.04
Chin-ups on gymnastic rings	.03
The Sargent jump	.18
A one hundred-yard shuttle run	-.15
Multiple fitness index	.10
The General Aptitude Test Battery	**Pearson r**
General learning ability	.30
Verbal aptitude	.25
Numerical aptitude	.31
Spatial aptitude	.17
Motor coordination	.30
Manual dexterity	.12
The Personality Inventory	**Pearson r**
Personnel reaction blank	-.33

Correlations significant at the .05 and .01 levels (N = 212) are .14 and .18, respectively. The negative correlation for the shuttle run was due to a time score, and the negative correlation for the personnel reaction blank was caused by a reversal of the scoring key from the usual order. The multiple fitness index was an unweighted composite of the five preceding tests. It was necessary, due to insufficient administration time, to delete sections of the GATB which yielded scores for form perception, clerical aptitude, and finger dexterity. Only one of the 10 predictors from the application record correlated high enough with work efficiency ratings to be practically important. This was years of education, which yielded a correlation of .34, the highest for any single predictor. A multiple correlation of .43, with a .06 standard error, was obtained, based on the personnel reaction blank, numerical aptitude, and motor coordination.

This study lends itself to a variety of interpretations. On the positive side, it is interesting to note that the personality inventory and years of education exceed slightly all of the GATB factor scores in predictive power. The personnel reaction blank also contributed significantly in the multiple correlation studies. On the other hand, the criteria were so general that the various correlations cannot be meaningfully compared. The personality inventory and years of education may have been better predictors of success in a number of jobs combined (as in this study), whereas aptitude scores might predict better in a specific job. It is interesting that significant correlations were obtained between the predictors and such a general criterion of work efficiency ratings regardless of which or how many jobs the worker had performed. Studies of success in specific jobs, or closely related jobs, are far more common in the literature, and probably are to be preferred as a basis for the selection--and certainly for the placement--of new workers.

The McEwen Study (1962)

This was a follow-up study of former students in programs of vocational agriculture in Idaho secondary schools. The data were obtained from instructors, rather than the students themselves, in 23 selected schools. The instructor had been employed in each case for the past 10 years. The names of 228 students composed the sample, all of whom were high school seniors during the 1955-56 school year. Returned questionnaires from the instructors accounted for 218, or 95.6 per cent, of the 228 former students.

The primary purpose of the study was to determine what percentage of the former students were engaged in farming or in an agriculturally related occupation. The results revealed that 33.2 per cent were engaged in farming, 9.7 per cent in a field related to agriculture, 39.5 per cent were employed in a field not related to agriculture, and 17.6 per cent were enrolled in college. At one time or another, one hundred of the students had enrolled in college, and 45 per cent of these had selected some phase of agriculture as a course of study. The percentage engaged in farming was higher than similar percentages obtained in certain studies completed elsewhere in the United States.

Several additional findings were presented, as McEwen asked several questions pertaining to the experiences of the former students. One portion of the thesis, of particular interest in this report, pertained to factors influencing the students to engage in farming. These factors were: highly interested in farming as an occupation, vocational agriculture training received while attending high school, father established in farming, secured partnership with parents, influenced by vocational agriculture teacher, development of an outstanding supervised farming program while in high school, inherited a farm, and acquired a farm by financial loan. These factors might be classified under the broad heading of either biographical information or socio-economic characteristics. Factors of this type are important and have proven their utility many times in studies relating to vocational guidance, selection, and placement.

It seems evident, however, that the findings can be accepted as only suggestive. The most obvious point of concern is that vocational agriculture instructors were asked to supply the information about the students. This method did result in a high percentage of returns, but it is questionable if agricultural instructors have this much detailed information about their former students. There also is a tendency for bias; even if the instructors try to be as objective as possible, they may tend to slant their replies in a direction which favors their local program. The study, typical of many follow-ups, attempted to integrate a number of specific responses into a meaningful picture of the post-graduation activities of the former students. A better design and some refined statistical analyses would have improved the study.

The McHugh Study (1967)

McHugh conducted a longitudinal study of 146 trainees in classes organized under the Manpower Development and Training Act (MDTA). The training was conducted in the area served by the Twin Falls local office of the Department of Employment, State of Idaho. The economy of the Twin Falls area is basically agricultural, and the occupational opportunities are primarily affiliated with farming or farm-related industries. The group of 146 trainees consisted of 99 males and 47 females. They were divided further into 48 persons who were trained in clerical and sales, 40 in farming or related, 15 in machine trades, 28 in structural work, and 15 in miscellaneous occupations.

The trainee characteristics selected as independent variables were age, educational attainment, number of dependents, number of prior jobs, years of gainful employment, occupational norms as determined by the General Aptitude Test Battery (GATB), and general learning ability determined from the GATB. The foregoing data were obtained prior to the beginning of training. The trainees were followed up through their training programs and for 90 days after their training was completed. Two groups were formed. Successful trainees were defined as those who successfully completed a prescribed training course as determined by training institutions and who 90 days after completion were employed in the occupation for which training was received. Unsuccessful trainees were those who withdrew before completion or were not employed 90 days after completion in the occupation for which training was received.

McHugh stated as a general hypothesis that there are factors of age, educational attainment, number of dependents, number of prior jobs, years of gainful employment, occupational norms as determined by the GATB, and general learning ability as measured by the GATB are independent variables known by the Employment Service and that these factors can be used as variables in determining discriminant functions for the selection of individuals for training programs in short term occupational choice. McHugh had to reject this hypothesis, as none of the resulting values were significant.

This study is a good example of a research project which investigated the training process over a period of time. Relationships between data collected before training begins and later criterion variables, if significant, have potential use in vocational guidance. McHugh did not discuss, however, the possibility that these trainees may have been subjected to a certain amount of selection based on one or more of the independent variables he studied. The extent to which such selection factors may have operated could be expected to influence group comparisons once the training is completed. Furthermore, since no significant discriminations were found, more studies obviously are needed to determine if this general research design will yield information of practical value for vocational guidance, selection, and placement.

The McRae Study (1968)

McRae conducted a follow-up study of graduates from three vocational programs in the Trade and Technical School at Idaho State University, Pocatello. The study was quite similar to the follow-up conducted by Gentry (see above, pages 52-53), and also was influenced by the Fifield-Watson study (see above, pages 49-50). McRae mailed one hundred questionnaires and received 61 usable returns. The programs and numbers of graduates in each were: auto mechanics, 20; drafting and design, 31; and 49 in electronics.

This seems to be a useful study, although it is limited in scope. A majority of the graduates reported employment in occupations related to their training; this was true of less than half of those who responded in the Gentry study. McRae's study did not include too much statistical manipulation of data, nor very many interpretations. The tendency was to tally the responses, compute certain percentages, and leave it to the reader to extract additional meaning from the results. School officials planning follow-up studies should examine this study for ideas on the types of questions that might be directed to former students. Other studies, including some cited in this chapter, would have more to offer regarding analyses and interpretations of the results obtained.

The Meyerhoeffer Study (1967)

Students enrolled in the Area Vocational-Technical School, College of Southern Idaho, Twin Falls, were the subjects of a validation study of tests and a derived measure of motivation. The tests were the General Aptitude Test Battery (GATB) and the Gates Reading Survey. Motivation was measured by class attendance. The samples were composed of all students enrolled during the first semester of the 1966-67 academic year. Approximately 90 per cent were males, and about one-fourth were married.

The students were ranked in their major fields in terms of the Occupational Aptitude Profile on the GATB, the Gates Reading Survey, first semester grade point averages, and first semester attendance. The following rho coefficients were obtained:

(Note: See first entry on the next page)

Since the number of cases differed for motivation, the following results are presented separately:

(Note: See second entry on the next page)

Class	N	GATB & Class Rank	Gates & Class Rank	GATB & Gates
Agricultural equipment technology	10	.015	.288	.649*
Automotive body repair I	11	.214	.759**	.224
Automotove body repair II	11	.025	.080	.591
Automotive mechanics I	16	.165	.451	-.120
Automotive mechanics II	14	-.260	.570*	.270
Law enforcement	10	.164	.509	.521
Mid-management	10	.515	.770*	.441
Radio-television technology	11	-.060	-.225	.600
Welding technology	14	-.410	.446	-.709**

Significance: * = 5%; ** = 1%

Class	N	Attendance & Class Rank
Agricultural equipment technology	11	.580
Automotive body repair I	19	.767**
Automotive body repair II	14	.729**
Automotive mechanics I	18	.545*
Automotive mechanics II	14	.691**
Law enforcement	10	.818**
Mid-management	17	.679**
Radio-television technology	11	.743**
Welding technology	14	.444

Significance: * = 5%; ** = 1%

It is very interesting to note that the Gates correlated higher (disregarding signs) with class rank based on grades than did the GATB in all of the nine vocational major fields. Since the GATB and Gates are not highly correlated, the two instruments might be combined to advantage in a multiple prediction study. The measure of motivation does not seem to be a very logical or useful predictor, but the results obtained indicate that it correlated with class rank higher than either the Gates or the GATB. Further study would be needed to determine how all of these relationships would hold, especially with larger numbers of cases. The Gates and the GATB have several parts, and correlations would be easier to interpret if they were based on individual scores. The derived measure of motivation should be explored further; it may reflect either teacher reactions to absences, student characteristics important for achievement, or a combination of these and/or other factors. Meyerhoeffer included several recommendations, and a suggestion that the GATB was not very useful in selection. The latter statement cannot be supported with such small samples, but the results indicate that validation studies should include other instruments such as reading tests in addition to the GATB.

The Nybroten Study (1964)

This was a study of the career plans of high school seniors in 10 counties of southern Idaho. An added feature was the presentation of labor force data for both Idaho and the nation. Several comparisons were made between the career choices of the seniors and the structure of the local, state and national labor force. Data concerning the number and distribution of the labor force in the 10 counties were collected by questionnaire from employing establishments. The Department of Employment (formerly the Employment Security Agency) directed these studies in each county, and 91,429 workers were represented. The Idaho State Board for Vocational Education used a simple questionnaire to obtain the occupational plans of 3,654 high school seniors.

Certain national and local labor force trends were identified, such as the decline in agricultural employment and the increase in the employment of women. Idaho was shown to have a higher percentage of the labor force who were high school graduates, but a smaller percentage of these graduated from college than elsewhere in the nation. Compared with anticipated openings in their own areas, several discrepancies were noted for the high school seniors. For example, more seniors planned professional careers than there were potential openings, whereas an undersupply was indicated in other areas.

This study resembles that of Brenna (see above, pages 42-43), and the two studies may have had certain data in common. Therefore, most of the comments regarding Brenna's study are applicable here. The writer has had experience with the instruments used by both the Department of Employment and the Idaho State Board for Vocational Education. Both agencies used fairly simple instruments to gather their data. The resulting data sometimes were difficult to interpret, and comparisons could be made only on the basis of fairly broad categories. Even so, Nybroten's study provides some useful information, and reveals a need for (1) more career guidance and dissemination of occupational information, and (2) follow-up studies to determine the relationships between stated career preferences and actual employment.

The Olson Study (1966)

A follow-up study was conducted of 410 graduates of the class of 1959-60 at the Idaho Falls High School. Mail questionnaires were used to contact the graduates, 223 of whom (112 males and 111 females) responded for a return of 54 per cent. The study had two main purposes: (1) to investigate the present status of former high school graduates, and (2) to secure their evaluations and recommendations pertaining to the effectiveness of the school's educational program. Data also were obtained for analysis from student transcripts.

A fantastic amount of data was collected in the study. An effort was made to obtain information regarding geographic location, additional education, family life, employment, earning power, and other areas of activity in the lives of the students. The graduates were asked to evaluate the high school administration, faculty, discipline, physical features, and other aspects of the high school's educational program.

The results of the study were many and varied. Several data were compared and frequencies, trends, and other characteristics were noted. Examples only of the findings can be presented here. Two rather negative results were: (1) the respondents received very little help from the high school in securing employment upon graduation, and (2) slightly more than one-third of the respondents in their written comments stated that they were not adequately stimulated to learn. Examples of two findings on the positive side were: (1) the majority of the respondents appraised the high school administration from good to excellent in most areas, and (2) 200 of the 223 respondents felt that the high school adequately recognized the students in planning school activities.

The instrument developed for this study is an excellent example of a follow-up questionnaire. It was used in the Fifield-Watson study (see above, pages 49-50) as a guide in the construction of their follow-up instrument. Olson also served as a consultant to Fifield and Watson during the initial planning stages of their study. A major weakness in Olson's study was the proliferation of various findings and relationships with few summary statements or efforts to interpret the data. The reader tends to lose track of earlier material, and it is difficult to grasp the overall significance of the findings. There also was little use of statistical treatment, as many data were presented in frequency form and not even converted into percentages for comparison. Finally, certain data which might serve as criteria, such as employment experiences and college major fields, were not well developed.

The Rutledge Study (1967)

Dropouts from the Caldwell, Idaho school system during the school year 1964-65, including the summer of 1964, were the subjects of this study. An effort was made to determine the extent of the problem (percentage of dropouts) and factors leading up to and characterizing dropout behavior. Data were obtained from four sources: (1) a questionnaire sent to the dropouts, (2) a form completed by the teachers of the dropouts, (3) the dropout's school file, and (4) some interviews with the dropout, his family, or others who knew him. This is a good example of a study in which data were gathered from a variety of sources. School counselors contributed interview information obtained from several dropouts at the time they left school. There were 42 dropouts during the year, which was a rate of 3.2 per cent. Only about half of these returned questionnaires, but fairly complete data were available from other sources.

The results were analyzed fairly extensively, and several comparisons were made with the results obtained in similar studies in Idaho and elsewhere. This also makes the study an exemplary one in regard to the types of data collected and the comparisons which followed. Individual case studies were included to give more meaning to the general findings. One salient finding was that 9 of the 42 dropouts were trying to finish their schooling. Certain distinguishing characteristics of the dropouts were: (1) average levels and ranges of IQ and achievement scores, (2) predominance of girls who had quit to marry, (3) a majority of over-age students, and (4) the high percentage of broken homes of the dropouts. The small number of cases indicates that the findings should be interpreted cautiously.

The Scharf Study (1966)

Since counselors are perhaps the most logical users of instruments for vocational guidance, selection, and placement, the selection of counselors themselves is a very timely and worthwhile topic for study and research. Scharf undertook to develop a counselor selection scale which would differentiate counselors as a group from persons in other educational positions. The rationale upon which the study was predicated was that counselors possess qualities which other educators either do not possess, or possess to some degree less than counselors. It was assumed further that this difference in counselors would express itself as a kind of "guidance-mindedness" and that this quality was amenable to measurement. Scharf patterned his study after an earlier study completed in 1947 by his major professor (Dr. Eugene Giles, formerly Professor of Psychology, University of Idaho) as a doctoral thesis at the University of Washington, Seattle.

The instruments were assembled and administered to samples of educators in Idaho, Montana, Nevada, Oregon, Washington, and Wyoming. The measurement battery was composed of both commercially available instruments and others privately developed. The components were the Allport-Vernon-Lindzey Study of Values; an interest inventory developed by Giles in his earlier study; the California Test of Interests and Activities, part of the California Test of Personality; and a Word Preference Test developed by the investigator. A preliminary administration was conducted on 42 educators attending a summer session at the University of Idaho. The purpose was to refine the instrument. The final sample contained educators holding seven different types of educational positions, under the three main titles of Guidance, Administration, and Teaching. The battery was mailed to a potential sample of 210, but only 82, or 39 per cent, were returned.

Individual item responses were analyzed and compared for the seven groups of educators, using the chi-square technique. It was found that 20 of the 96 questions composing the final battery were answered differently by the groups studied to yield significant chi-squares beyond the .05 level. The counselors were shown to be significantly different on 10 of the 20 questions compared with more than 2 of the other 7 groups. The remaining 10 questions differentiated the counselors, but less clearly.

Scharf has shown that counselors can be differentiated from persons in other educational specialties. His samples were very small, and this he readily admits. It is very interesting to note that the same year Scharf completed his study, the Kuder Occupational Interest Survey, Form DD was published (Kuder, 1966, pp. 44-46) which included scales for both men and women for Counselor, High School. The men's scale was based on 200 cases, and the women's on 250 cases, all of whom were members of the American Personnel and Guidance Association. Scharf thus has a commercially developed instrument with which he can compare the results of any further studies he may complete with his battery. He may be able to make a significant contribution to counselor selection, but much more research will be needed than is reflected in his doctoral thesis.

The Sorensen Study (1967)

The major purpose of this study was to determine the present occupational status of vocational agriculture graduates during the period 1956 through 1965. High school graduates during 1956, 1957, 1959, 1961, 1963, and 1965 were selected for study. During these years, 69 vocational agriculture departments existed in Idaho. A simple questionnaire was prepared and sent, not to the students, but to the school departments. A total of 61 out of 69 responded, which Sorensen interpreted as an 88 per cent return. The results pertained to 2,617 students, but there is no indication of the total student population.

The occupations reported for the graduates were assigned to 10 broad categories, with the following result:

Occupation	Number of Persons	Percentage
Full-time farming	422	16.13
Part-time farming	38	1.45
Agri-business	403	15.39
Non-farm	627	23.96
College	573	21.89
Armed forces	251	9.60
Missionaries	70	2.67
Deceased	31	1.19
Unemployed	16	0.61
Unknown	186	7.11
Totals	2,617	100.00

The above figures and percentages were presented in Sorensen's Table 1 (1967, p. 21).

The writer was not very pleased with this study for several reasons. Sorensen states that about 58 per cent of the vocational graduates were engaged in farming or other agricultural pursuits. This is inconsistent with the above figures; the first three entries total 33 per cent, not 58 per cent. The 33 per cent figure is comparable to the results of another study cited in this chapter (see McEwen, pages 59-60). Only the highly selected group of State Farmer recipients studied by Holmquist (see above, page 53) had more than half engaged in the broad field of agriculture (48 per cent in farming and another 18 per cent in agriculturally-related jobs). The decision to send questionnaires to vocational agriculture departments is questionable. The accuracy of the results obtained were dependent upon how closely the departments had followed former graduates. Sorensen also made several recommendations for curriculum revisions in his discussion section which do not seem to be derived from his data. This study had some potential for yielding useful information, but certain of the assumptions made and the way the data were treated and presented render it unsuitable for any serious conclusions.

The Stewart Study (1968)

The major part of this study was conducted in the state of California. Students from one Hawaiian community college and from five post-secondary area vocational schools in Idaho also participated in the study. The latter portion of the study was sponsored by the staff of the Idaho State Occupational Research Unit, University of Idaho. The investigator and his associates at the University of California, Berkeley developed a new instrument to measure interests, named the Interest Assessment Scales (IAS), which yielded measures on eight separate test scales: adventure, order, influencing others, nurturance, concrete means, written expression, abstract ideas, and aesthetic. Another instrument, the Omnibus Personality Inventory (OPI), also was used in the study, and yielded the following seven measures: autonomy, complexity, estheticism, impulse expression, social introversion, thinking introversion, and theoretical orientation. The two instruments were administered to California students in 43 vocational curricula, graduates of a number of the vocational curricula who had entered a job related to their training, and a small group of apprentices. Additional samples, as indicated above, were obtained in Hawaii and Idaho. Data analyses were performed using two powerful statistical methods, namely, stepwise multiple discriminant analysis and multivariate analysis of variance.

Since the Idaho portion of the study is of most interest here, some additional detail is needed. It was necessary in Idaho to administer a different form of the OPI which yielded scores on 14 instead of 7 scales. The additional scales were as follows: religious orientation, personal integration, anxiety level, altruism, practical outlook, masculinity-femininity, and response bias. The OPI was administered to 719 Idaho subjects; due to losses encountered during scoring and for other reasons, only 463 subjects had complete results on the IAS. The analyses involving performance on the OPI were based on 17 vocational curricula, whereas the number was reduced to 14 for the IAS analyses.

The instruments both discriminated significantly between vocational curricular groups in all three states. The IAS was found superior in this respect to the OPI. The IAS results in Idaho, subjected to a stepwise multiple discriminant analysis, correctly classified from 0 to 50 per cent of the subjects into their respective curricular groups. Overall, 24 per cent of the subjects were correctly classified. A multivariate analysis of variance of the IAS results indicated that the curricular groups differed in the factor structure underlying their scores.

The results also revealed significant discriminations between the vocational curricular groups. The OPI results were similarly analyzed. The discriminant analyses correctly classified from 0 to 75 per cent of the subjects, but overall only 18 per cent were correctly classified. The multivariate analyses of variance were roughly similar to the results obtained for the IAS. The percentages correctly classified in Idaho were lower than in the other two states (California and Hawaii).

This study was designed and executed better than any similar study that has come to the attention of the writer. It is a prime example of a study concerned with group membership, as opposed to predictions of success within a group (Cf. Chapter I, pages 5-7). Perhaps the relatively low percentages of subjects correctly classified in the Idaho portion of the study reflects the rapid growth of vocational education in this state in recent years. Both the course content and admission procedures in various vocational curricula may have been more subject to change and thus less likely to yield groups of subjects with similar interest and personality patterns. Perhaps the main weakness of this study is its complexity. It takes a tremendous amount of research and statistical knowledge just to read the report with understanding. Moreover, the results obtained from the statistical analyses are not readily converted into useful predictions as is possible in a regression equation derived from correlation studies. The IAS, particularly, is very difficult to score, and is still in the experimental stage. Although the study may be considered a landmark in research concerned with vocational students, much additional development is needed before similar results can be put to common use in vocational counseling, selection, and placement.

The Williamson Study (1964)

A variety of instruments has been developed to measure interests. Williamson tried a new approach, and tested the hypothesis that an individual's interests can be determined by his selection of definitions of words which relate to his interests. It was postulated that individuals would tend to select a specific definition for a word in preference to other possible definitions depending upon his unique interests. A special instrument was developed by selecting 765 words from the 8,000-word level or below in Thorndike's list of the 20,000 most frequently used words. Four definitions were written for each word from the Funk and Wagnalls dictionary. The testees were instructed to select only one definition per word with which they were most familiar or used the most.

The completed instrument was administered to four criterion groups specifically chosen for the study to be relatively homogeneous as well as different from each other. These were: 20 agricultural instructors in public high schools, 14 electrical and chemical engineers, 20 elementary teachers in public schools, and 12 senior law students at the University of Idaho. The four criterion groups composed six paired groups, each of which was compared on their responses to each word. A binomial approach was used which compared the percentage responses to each of the four alternative definitions. The results indicated that 125 of the 765 words discriminated significantly one or more of the criterion groups from at least one other group. The hypothesis thus was sustained. This was a creative and fairly sophisticated study. The real test of this approach is whether it will prove superior to instruments like the Strong or Kuder, or if it will yield useful information not presently available from other instruments.

The Witty Study (1962)

This study grew out of a desire, on the part of the author, to increase student motivation in English classes in high school. It was felt that this could be achieved by finding concrete evidence of the need for English competencies in various institutions of higher learning and also in local business firms employing workers at the sub-baccalaureate level. That portion of the study dealing with higher educational institutions was based primarily on an analysis of college catalogs for a 20 per cent sample taken from the 1956 Educational Directory. The results left little doubt that English is a key subject not only for college entrance but also for progress through a variety of college fields of study. For example, no matter what else he studies, a student could expect English to be a part of at least the first two years of his college career.

The second major portion of the study was concerned with local business establishments in Boise, Nampa, and Caldwell and vicinity. A questionnaire was developed which posed such questions as numbers employed, hiring policies, and several questions pertaining specifically to needs and requirements for English competencies. The questionnaires were mailed to 136 business firms obtained from the Employment Security Agency (now the Department of Employment). These firms employed 75 per cent of the workers in the Boise valley. Questionnaires were returned by 112 firms, or 82.4 per cent, of which 111 were used for analysis.

The results tended to support the importance of English in work performance. Two salient findings were (1) one businessman in four believed that English grades were the most significant of all grades in measuring potential ability to handle the job, and (2) a student seeking employment in these establishments could anticipate that one businessman out of two would expect him or her to have completed four years of English. Witty points out, however, that such findings are difficult to interpret, as hiring requirements and procedures did not seem to be designed to check very closely on achievement or proficiency in English.

A study of this type can yield useful information in that some effort was made to get documented evidence of the need for English competence both at the college level and in the "ordinary" world of work. Too often courses are taught with little more than educated guesses as to their importance in later work activities. The questionnaire sent to employers was fairly simple, but might be used as a starting point toward something better. Witty herself found that the general questions asked, which pertained to many different jobs, caused the respondents some difficulty in completing the form. Future studies might focus on individual jobs, and personal interviews might be used to advantage. The employees who do the work also should have an opportunity to identify educational courses and experiences that helped prepare them for their jobs. Witty assumed that having some documented evidence of the need for English would result in increased student motivation. This might work with some students, but the manner of presentation in the classroom, and other factors, probably would be the key to motivating several other students.

The Young Study (1965)

A single study among those reviewed was concerned with development of achievement tests for electronics. Young developed his tests using subjects in California, but the work was completed as a thesis at an Idaho institution. The methods used, and the instruments developed, might be used in local studies. Multiple choice tests were developed for three subject matter areas in electronics: basic electricity, shop practices, and vacuum tube theory. The tests were administered to 51 students entering the electronics program at California State College, Long Beach, and to 113 fourth semester students in five junior colleges. Refined methods of item analysis were applied to the instruments. There also were considerations of adequacy of item sampling, reliability, external and internal validity, and difficulty. However, no validation studies were conducted. The principle use suggested was as additional data for counseling entering students.

These certainly are not finished instruments, as Young is quick to point out. He also recommends validation studies using a predictive validation design. A project of this type in Idaho area vocational schools which offer electronics programs might be worthwhile. Young used several sources to develop his instruments, such as radio operator license exams, textbooks, and teacher-made tests. The resulting instruments might be useful as achievement and evaluation instruments. On the other hand, the big question is whether all of the necessary effort and expense would produce a set of measuring instruments that would improve the predictions which could be obtained from existing commercially available tests. There also might be a tendency to develop local achievement type tests which would need frequent revision to keep abreast of new developments. The resulting expense might outweigh the usefulness of the instruments.

DISCUSSION, SUMMARY, AND CONCLUSIONS

Discussion

Certain general characteristics of the 31 studies presented in the preceding section are of interest. The vast majority were research projects conducted as part of the requirements for a graduate college degree. There were 3 masters papers, 19 masters theses, and 4 doctoral theses; the remaining 5 studies were sponsored by private individuals or by schools or other organizations. The decision to include studies using some type of privately developed questionnaires increased considerably the number of studies cited. There were 16 studies based on questionnaire results, and an additional 5 which had some type of privately developed testing instrument. The remaining 10 studies either utilized a commercially developed testing instrument, or this in combination with a privately developed instrument.

Graduate research projects related to degree programs have several obvious advantages. One example is that they tend to be good sources of research findings that reflect "ground breaking" in new field of inquiry. A second example is that most research papers and theses reflect a respectable amount of research expertise; care in the analysis, presentation, and interpretation of data; and a written report that presents a rounded picture of what was accomplished, and the results obtained. There are, of course, too many exceptions, which leads to a discussion of some important negatives. First, graduate student research tends to be individual. One study thus is difficult to compare with another. True, there were several studies in the preceding section concerned with such broad topics or fields as agriculture, validation of psychological tests, and student follow-up, but the designs, samples, etc. differ to the extent that the results usually are not comparable or additive from one study to another. Second, graduate students serve several masters during their research in the form of the major professor and committee members. This affects the kinds of data and problems which may be selected, and these may not provide direct answers to immediate practical problems. More private research, or research by organizations and agencies, directed toward problems common to Idaho, is definitely needed.

Despite their differences in design, the four studies which utilized the General Aptitude Test Battery (GATB) yielded results which are important to the overall purpose of the present report. The study by Burke (see pages 44-46) was most similar to those conducted by the writer and reported in Chapter III. Burke's results are more similar to those obtained for electronics students at Boise State College than at Idaho State University (see Tables V and VI, pages 30 and 34, respectively). Once again, there are sufficiently large differences in the predictive value of individual aptitude factors to justify separate validation studies. The correlations obtained by Burke are not high, and the multiple correlation of .411 is below the .45 level adopted by the writer as a minimum for practical prediction. The study by McHugh (see pages 60-61) was based on a division of the trainees into successful and unsuccessful groups. Different criteria, and the use of correlation techniques rather than the discriminant function, might produce a better and more productive research design.

The remaining two studies by Loudermilk (see pages 57-59) and Meyerhoeffer (see pages 61-62) are of interest because they included other instruments in addition to the GATB. Combined with the results obtained at Boise State College (see Table IV, page 29), it is evident that other predictors should be combined with the GATB in research studies. Reading tests, personality inventories, measures of motivation, biographical information, and other measures may account for additional criterion variance which is not predicted by the GATB factors. Another important consideration is that the GATB is a factored aptitude battery. Instruments which are relatively pure factorially tend to be less useful as predictors of practical criteria than more complex tests (Michael, 1958; Loudermilk, 1966). The only logical conclusion is that experimental test batteries should contain a wide variety of instruments and not just the GATB or a similar battery of aptitude tests.

Summary

This chapter has presented resumes and critical appraisals of 31 separate research studies in Idaho. The vast majority of the studies reviewed were completed as part of the degree requirements by graduate students in institutions of higher learning. Studies using questionnaires and testing instruments privately developed accounted for more than half of those reviewed. A need was identified for research to be completed by private individuals and organizations which is not influenced by the rigid requirements of a graduate research paper or thesis. A few studies were found which revealed that the General Aptitude Test Battery needs to be supplemented by other types of tests which likely will improve predictions. Generally speaking, the studies cited were exemplary, but several weaknesses were found which should be corrected to improve future research studies.

Conclusions

At the beginning of this chapter, the writer made it clear that there would be an intentional emphasis on criticisms of the studies reviewed. This also will be true of the conclusions which follow, at least in regard to space allocations. Perhaps the reader should give about as much weight to the first conclusion alone as he does to all of the rest which follow. Once again it may be stated that the writer's intention is to promote better research through a critical appraisal of the work completed to date.

Based on the contents of this chapter, and appropriate references to earlier chapters, the following conclusions may be stated:

--The research studies reviewed generally reflect considerable research expertise, care in design and analysis, and adequate attention to tabular, graphic, and written presentations of the results.

--The vast majority of research studies in Idaho relating to vocational guidance, selection, and placement have been graduate research projects in institutions of higher learning. While thesis research has definite values, there is a need for studies by individuals, agencies, and organizations which can focus more on existing practical problems.

--The General Aptitude Test Battery has potential value as an instrument for vocational guidance, selection, and placement. It may, however, be equaled or exceeded by other instruments in a number of studies, and therefore should not be the sole component of an experimental battery of tests.

--None of the research studies reviewed were entirely free of certain weaknesses which could be improved upon in future studies. This writer concluded that the following were sufficiently important to deserve a separate listing under this general conclusion (no order of imporance or frequency is intended).

a. Several instruments were developed privately by different investigators and were used as the means to collect data. Perhaps the best justification for this procedure is that no instrument available can be used to fulfill the objectives of the study. Some studies, however, included the development of an instrument which resembled commercially available instruments, or were intended as "new" measures in an area already characterized by significant quantities of research. The latter type of study has made no real contribution to research knowledge unless it can be demonstrated that the instrument is at least as good as, or better than, instruments already available.

b. Verbal responses only to several instruments, or different parts of a single instrument, constitute the only data in much of the research completed. Data derived from actual behavior are preferable as bases from which conclusions may be drawn.

c. Instruments which are used to collect data relatively limited in scope are sometimes considered as measures of concepts which are so complex that the instrument really can yield only a few correlates of the concept. A good case in point is the "self concept" in vocational choice theory. Since the vocational self concept is interpreted by many writers as virtually synonymous with the total personality, it is highly presumptuous to attempt to measure the self concept with anything less extensive than a complete personality assessment battery.

d. Too many studies were characterized by "raw empiricism." That is, results were presented willy nilly with no attempt to relate them to some theory, or even some logical design or framework which would order the results and facilitate interpretation. A related weakness in some studies was the tendency merely to present tables of data without adequate discussion and interpretation of the results.

e. There was a definite need for more refined statistical analyses. Most of the questionnaire studies were characterized by nothing more profound than the conversion of frequencies into percentages. Certain simple significance tests, and, where indicated, more complex statistical analyses, would have improved several studies.

f. Small samples were all too frequent in the studies cited. It would seem that, the smaller the sample, the more certain the researcher should be that the cases chosen are representative of the population being studied. Return rates on questionnaires also are pertinent here. Results from studies with returns below 50 per cent should not be reported unless it can be demonstrated that the respondents are not a biased group.

g. Several of the privately developed instruments contained "open-ended" questions. This is desirable in many ways, as it affords the respondent an opportunity to include pertinent information that might be omitted or obscured in a more structured instrument. It was evident in several studies, however, that a lack of structure in the instrument yielded almost a hodge-podge of responses which were difficult to interpret. Such data can be handled in a meaningful fashion, but it requires a carefully constructed design or rationale for their treatment and interpretation.

h. The test validation studies were not always free from prior student selection based on the same measures. Selection testing tends to rule out certain cases who otherwise might enroll in a vocational study program. This in turn tends to restrict the range of performance on the testing instrument and thus attenuate the correlation between test performance and training criteria.

i. A real need exists for criterion selection and development. Grade point averages and graduation *versus* non-graduation from educational programs have been almost too common in research studies. Career patterns, job satisfaction, interpersonal skill development, and occupational level are a few examples of criteria which might be developed and used in future studies.

j. More attention should be given to the basis for rejection of an experimental hypothesis. This is more involved than the familiar levels of significance adopted for hypothesis testing. For example, hypotheses in null form may be tested by significance tests for two groups compared on, say, 30 variables. A certain small number of significant differences might be expected to occur by chance. It seems evident that a null hypothesis could be rejected more convincingly if all 30 variables showed significant differences as compared with rejection based on a half dozen or less significant differences. The basis for rejection of the null hypothesis in such cases should be carefully worked out and stated in advance. Different statistical techniques also might be applied which are more suitable than, say, a *t* test of significance.

k. A few of the designs could be "tightened up" to advantage. Certain questions posed at the beginning of a study, and particularly certain theoretical propositions, were so general that considerable ambiguity was present in the design. General or ambiguous questions should be avoided, and theoretical propositions also if they cannot be reduced to reasonably precise statements which are amenable to testing and interpretation.

l. Statistical analyses should be relevant to the overall research design. A few studies were predicated on statistical analyses so complex that one suspects they were chosen to impress certain colleagues rather than to treat the available data. Moreover, statistics should be no more complex than are warranted by the nature and precision of the data to which they are applied.

m. Data collected for research studies should be as accurate as possible. This mainly involves getting as near as possible to the primary source. For example, student data should be collected from the students, not from their instructors, unless it can be demonstrated that the latter have the information sought or if the design specifically calls for information which logically should be supplied by instructional personnel. A related concern is that questions developed to elicit certain data should not be "slanted" toward a group or conceptual area being studied unless this is a logical part of the overall design.

n. Research performed on a selected sample should be repeated, if possible, on other samples. A sample of school dropouts, for example, may have numerous characteristics which are common to a high percentage of their members. Comparisons with "normal" groups, however, may reveal that several of the characteristics are common across several groups and thus of little value as indexes of group membership. This also includes the necessity of cross validation in test prediction studies.

o. Research studies need to be planned over a period of time which include plans and arrangements for the recording and collection of data. Too many studies were conceived after data had been collected and assembled in school, employment, or other records. Research studies based on data already collected (possibly excepting certain arrangements like formal data banks) often must be altered and/or curtailed to conform to the existing data.

CHAPTER V

NATIONAL DEVELOPMENTS IN VOCATIONAL APPRAISAL

At this point we have completed a review of research directly related to student appraisal in vocational-technical education programs in Idaho. The main task which remains, and which is completed in the next, and final, chapter is to attempt a synthesis of the material presented. There remains, however, the need for a brief review of certain major developments outside Idaho which have a bearing on further progress in the construction, validation, and use of instruments for vocational appraisal

This field is vast, needless to say, and the writer hesitated even to include the present chapter for fear that he could not really do justice to so vast a body of knowledge. The decision reached was to adopt a "spotlight technique" which would focus on a few significant studies, and reviews which summarize vast segments of the literature. Conclusions were drawn which seemed to this writer to be supported by research on the national scene, and which form an important backdrop for continued efforts in Idaho. The studies and reviews chosen divided logically into four main sections entitled Historical Perspective, Recent Studies and Reviews, Career Choice and Development Theory, and Computerized Information Systems.

HISTORICAL PERSPECTIVE

Test validation studies pertaining to various occupations began early in the present century. Dorcus and Jones (1950) reviewed this literature up to around 1950, and issued the results as the Handbook of Employee Selection. These authors set up several criteria to evaluate the studies reviewed. It is a sad commentary on the quality of reporting and research during this period that only 427 references of more than 2,100 examined met the criteria of minimum adequacy and thus were included in their Handbook.

Ghiselli (1949, 1955, 1966) also became interested in a review of occupational validities for a wide variety of tests and similar instruments. Whereas Dorcus and Jones were content with the presentation of several individual studies, Ghiselli performed several integrations and analyses of his reviews. For example, his division of the studies into those using training criteria as opposed to proficiency criteria, the presentation of ranges of validity coefficients in different studies with the same type of instrument, and classification according to the Dictionary of Occupational Titles make the results much more meaningful and useful for research and practical applications.

Patterson (1956) performed a useful service when he surveyed the literature for tests which predicted success in trade or vocational school courses. He presented the results of each study in brief narrative form, and followed this with a summary statement of the types of tests which should compose a vocational prediction battery.

Super (1949) and Super and Crites (1962) have reviewed the literature on vocational appraisal and have presented the results in yet another way in book form. Different broad types of tests are discussed, along with the evidence and methods for their use in guidance, selection, and placement.

The testing field has grown to vast proportions from rather meager beginnings early in the present century. Anyone "worth his salt" as an administrator and user of tests is familiar with the _Mental Measurements Yearbooks_ (Buros, 1965; 1959, and other editions) and general books on testing by such noted authors as Cronbach (1960) and Anastasi (1968). Professional journals are another source of valuable information concerning tests and their use. Two excellent examples are the Validity Information Exchange published in _Personnel Psychology_, and the Validity Studies Section of _Educational and Psychological Measurement_.

The staff of the State Occupational Research Unit (1966) used most of the foregoing references, and others, in an earlier study to determine what types of tests should be included in an experimental testing battery for use in Idaho. The results of the review indicated that at least 10 broad types of tests might be used in test validation studies (1966, p. 29). Certain references were cited which revealed the necessity of local studies, rather than the use of tests validated in settings different from those in which they are to be used.

RECENT STUDIES AND REVIEWS

During the present decade, there has been a significant increase in research and publications pertaining to the world of work. Vocational guidance again has become fashionable. Perhaps the major incentive for this increased attention to the vocational area has been Federal legislation such as the Manpower Development and Training Act, the Vocational Education Act of 1963, and the 1968 Amendments to the latter Act. A few important studies which have appeared recently are cited in the paragraphs which follow.

The study by Stewart (1968), in which Idaho students were involved already has been cited in the preceding chapter (see pages 67-68). The interest and personality patterns of vocational students and trainees is an area which deserves careful study.

A review study by Prediger, _et al_. (1968) followed the tradition of earlier reviews such as those by Dorcus and Jones and Ghiselli (see preceding section). Prediger and his associates were interested in predictors of success in high school level vocational programs. Their review, covering the period 1954-1967, included over 1,200 correlations reported in 38 studies. The results were summarized for 10 predictor categories and 11 vocational program areas. The major findings were that (1) there was considerable variation in the level of correlation obtained for a given predictor within a given vocational area, (2) the predictability of success was much greater in some vocational areas than in others, (3) the predictability of grades was higher for girls than for boys, (4) the effectiveness of a given predictor varied from one vocational area to another, (5) the verbal IQ should not be used as the sole predictor in vocational studies, and (6) perceptual speed and accuracy tests and dexterity tests appeared to contribute very little to the prediction of success (excluding possible exceptions with specially constructed tests).

An extensive and very important study concerned specifically with vocational students is entitled Project Mini-Score. Conceived in 1965 and funded by a grant from the U.S. Office of Education, the project was designed to obtain comprehensive test and personal data on all applicants to the area vocational-technical schools of Minnesota. The data collecte will be validated on criteria for successful students who have completed training programs, and have then taken a job in the field, or a closely related field, for which trained. The test instruments included the General Aptitude Test Battery Form B (written portions only), the Minnesota Vocational Interest Inventory, the Minnesota Importance Questionnaire, the Vocational Development Inventory, and the Minnesota Scholastic Aptitude Test. The first major report of the project contains distributions of scores on the GATB for students who have graduated from various vocational programs (Pucel & Nelson, 1969).

Most readers of this report will be familiar with Project Talent, directed by Dr. John C. Flanagan and sponsored by the American Institute for Research and the University of Pittsburgh. The study was launched in March of 1960 when some 440,000 students in 1,353 public, private, and parochial secondary schools in all parts of the country were administered the Project Talent Tests (Flanagan, *et al.*, 1962, p. 14). During the annual convention of the American Vocational Association in Dallas, Texas in December of 1968, the writer was privileged to hear Dr. Flanagan (1968 present on a preliminary five-year follow-up of certain students tested in the twelfth grade. The means and sigmas on 15 selected tests for groups planning certain careers five years after graduation were presented in chart form. Only those persons actually employed in the job or who were training for the job were included in the statistical analyses. These preliminary data provide some important clues to the various attributes that may lead to success in a variety of work fields--both for men and women. Further analyses are being completed for each of more than two hundred occupational groups, which should provide important information for local guidance and research programs.

A study completed by Jex (1966) in our neighboring state of Utah is a good example of research which is needed for local programs. A series of tables were constructed to predict academic performance beyond high school. The predictions were presented in the form of grade point averages rather than expectancies. Jex found that the high school grade point average and an average achievement test score usually combined to yield multiple correlations with college scholarship in the range of .60 and into the .70 range. The high school grade point tended to be the best predictor. Little success was found in adding more or different kinds of predictors. Another important finding was that first-quarter college average was a better predictor of graduation from college than any other previous data from the high school record, thus suggesting that predictions could be placed on a continuing basis as the student progresses through school. Of most interest for the present report are predictions reported for various trade-technical curricula. Multiple correlations for the latter ranged from .50 to .73, all of which are above the minimum of .45 set by the writer of this report for practical predictions.

A study (Miller, 1968) involving high school students in two trade programs in Connecticut schools is an excellent example of how a test validation study might be organized and conducted. This was a cooperative effort between the state departments of education in both Connecticut and New York. Trade electricity and trade machine shop were the curricular areas selected for study, as they were offered in all of the 14 vocational-technical schools in Connecticut. The researcher began with the organization of a study team of professional consultants, and various committees assisted to keep the project moving ahead on schedule. The battery of tests consisted of the Flanagan Aptitude Classification Test, the Gordon Occupational Check List, the Stanford Achievement Test, the Primary Mental Abilities Test, and the Gordon Survey of Interpersonal Values. Criteria were developed specifically for the study in the form of a written test on technical information and a shop performance test. The resulting individual correlations were not high, but many were significant. Several multiple correlations were above the .45 level, indicating that useful practical results should be forthcoming when the study is completed. It should serve as a useful model for similar studies in other states.

CAREER CHOICE AND DEVELOPMENT THEORY

A very significant development during the last two decades is the appearance of a number of theories to account for the choice of a career and subsequent developments which form an individual career pattern or work history. One author (Borow, 1959) points out that, following World War II, there was real cause to doubt whether occupational psychology as a branch of behavioral science would survive. The fact that correlations were seldom higher than .60 (an observation which still seems to hold more than two decades later!) probably discouraged many researchers from expending additional efforts to improve predictions. Efforts to reduce the error variance present in occupational prediction studies, and other reasons, prompted the development of a number of theories to account for career choice and development.

The writer was very interested in this area of research more than a decade ago, and undertook a fairly extensive review of the literature early in the present decade (Loudermilk, 1961) which culminated in a research paper. The paper concluded with 11 critical evaluative comments pertaining to the literature reviewed. Since the writer believes these are still timely, they are included here:

1. Several of the theories (e.g. vocational interests, Roe's theory, need theory) seem too restricted or too loosely defined to be useful in all practical situations.

2. There are many areas in a broad theory of vocational development (e.g. Holland's) where we have not developed adequate measuring instruments.

3. Many of the instruments used in vocational development research are experimental instruments, and since the instruments vary so much from one study to another, results are difficult to compare.

4. The sizes of samples used in several of the research studies seem so inadequate that little, if any, support justly can be claimed for the findings reported.

5. Many of the empirical studies support hypotheses only on the basis of significant group differences. For individual counseling, such data might be meaningless.

6. The vast majority of all studies were made on students with vocational choice represented either by curricular choice, or merely by verbal expressions of students. Studies of workers choosing, and engaged in, actual occupations are needed.

7. Some of the studies have been supported by financial grants of one kind or another and are quite extensive; consequently, persons in doctoral research programs have difficulty in arranging studies which really seem worthwhile by comparison.

8. Many of the study techniques and statistical analyses may prove so expensive that they never will be available for common usage by counselors in the field.

9. There has been relatively little study of occupational choice at the lower occupational levels--for example, below the level of skilled occupations.

10. In many studies, when hypothesized results are not obtained, one almost gets tired of the profusion of apologies on the part of the researcher, as he tries to show why his study failed to yield expected results. In many, it would perhaps be more surprising if the results were significant.

11. Some theory builders, when studies of their theories fail to yield supporting results, will go to great pains to discredit the study. And yet they will not hesitate to point out any part of the same study which supports their theory.

The intervening years have improved the picture considerably as more research has accumulated, and several of the theorists have been able to rework their theoretical formulations and make them more amenable to experimental test. Recent developments have been ably presented and evaluated by Osipow (1968). He examined seven major viewpoints concerning the process of vocational decision making and subsequent career behavior. This writer gained several impressions from Osipow's work which bear directly on the main theme of the present report which is concerned with vocational guidance, selection, and placement. Regarding guidance, the theories are not sufficiently specific to enable a counselor to assist a client to make a specific choice. Only general predictions within fairly broad fields and levels of work seem possible at this stage of development. Their utility for selection and placement in specific training programs is limited further in that little attention is given to the role of aptitudes. The difficulties encountered in collecting good data, and the lack of good instruments, renders these theoretical approaches less effective than the more traditional forms of psychological testing. This writer believes that the theories give the counselor or user of tests a good conceptual framework, but until they are developed further and experimentally tested in a variety of situations, the use of more familiar testing instruments and research designs will yield the best results for practical use in Idaho's vocational-technical education system.

COMPUTERIZED INFORMATION SYSTEMS

The increased availability and use of electronic computers in recent years have resulted in their adaptation to programs of counseling and guidance. It is interesting that computerized vocational guidance was conceived and demonstrated more than 40 years ago by Hull (1925), but apparently his work was more than a quarter of a century ahead of its time. More recently, Cattell (1965) discussed a number of uses for the computer in addition to such standard functions as rapid test scoring, and the derivation of precise, easily storable descriptive records and standardized scores and profiles. A major concern is that the computer will be used to discover scientific psychological laws and derive predictions based upon these and others already established from previous research.

Loughary, et al. (1966) demonstrated how counseling processes and outcomes could be computerized for educational counseling with ninth grade students. Their study revealed that man-machine systems can be devised which have potential for improving the services rendered by professional counselors.

Computerized information systems specifically designed for vocational guidance have been developed. A prototype system was developed in the Department of Vocational Education at the Pennsylvania State University (Impellitteri, 1967). The computer is used to store up-to-date information about selected occupations, and to present the stored information to ninth grade students via typewriter printout, tape recordings, and slide projections. The presentation is selective; for example, the information presented to a student is based on his profile on the General Aptitude Test Battery. A more elaborate system is being developed at Harvard University (Tiedeman, et al., 1967; 1968).

Computerized vocational information currently is available to students in the Willowbrook High School, Villa Park, Illinois (Harris, 1968). The writer was privileged to hear Miss Harris present on this system at the National Conference of Research Coordinating Unit personnel during March, 1969 in Salt Lake City, Utah. The whole system is integrated with the guidance program of the school. Still another computerized information system in New York state is described by Minor, et al. (1969).

DISCUSSION, SUMMARY, AND CONCLUSIONS

Discussion

The studies and reviews cited in the first two sections of this chapter reveal a long history of research pertaining to the vocational area. This brings us to the current scene with a vast storehouse of knowledge concerning tests, statistical methods, and most of the other tools and techniques needed to use tests to improve our understanding, prediction, and to some extent the control of human behavior as expressed in various occupational pursuits. Perfection, or even near perfection, certainly has not been reached (Cf. earlier discussion, pp. 1-4). It seems evident, however, that we know enough about tests, and what they can and cannot do, that we can use them in ways which will result in definite improvements in the education and training of students for the world of work.

Theory construction and theoretical research pertaining to career choice and development have made a definite contribution to the fields of vocational psychology and vocational education. If nothing else, they have ordered an otherwise confusing variety of research findings into a logical conceptual framework. Another value is that their complexity and lack of specificity generate a healthy humility in many counselors who otherwise might tend to oversimplify this important aspect of the lives of their clients. This is a relatively young field of researcl activity. Consequently, definitive results supported by empirical investigations could not be expected at the present time. Future developments should be well worth watching, and research contributions would seem to be in order from those who have the necessary interests and competencies.

The electronic computer undoubtedly has tremendous potential as a tool to improve the breadth and quality of information which may be used in vocational guidance, selection, and placement. Two very important considerations come to mind, however, regarding the development and use of a computerized information system in Idaho. First, it is necessary to remember the watchword for all users of computers, namely GIGO, which stands for "garbage in, garbage out." It is the opinion of this writer, based on several years of research experience in Idaho, that the necessary information for a computer based information system is not available, and is unlikely to be for several years to come. The second consideration is related to the first, and pertains to the necessity of continual updating of the computerized information. In view of the statement just made regarding the lack of initial information for the system, the additional time and effort required for maintenance of the system indicates that the chances are slim indeed that Idaho can expect to develop a computerized information system in the foreseeable future.

Summary

A brief historical sketch has been presented of research with instruments having potential value for vocational guidance, selection, and placement. An impressive amount of research has been completed, and several publications present reviews, analyses, and integrations of previous research. A number of very recent studies have been concerned specifically with the vocational area, and indicate that vocational guidance and education are receiving increased attention by researchers. Career choice and development theories were considered, and computerized information systems were briefly discussed. Even though a respectable amount of research has been completed in the latter two areas, practical applications in Idaho were not deemed likely in the foreseeable future.

Conclusions

The contents of this chapter and the studies cited seem to support the following conclusions:

--An impressive body of knowledge pertaining to the use of instruments for vocational guidance, selection, and placement has been built up during the present century based on thousands of research studies. Principles pertaining to appraisal instruments and their use, and the types of instruments suitable for use in a variety of situations, are fairly well established.

--There remains the necessity of research with appraisal instruments in the specific setting in which they are to be used. Subsequent use should be tempered by an understanding of what the instruments can and cannot do.

--Due at least in part to recent Federal legislation, vocational guidance and vocational education are respectable fields in which to generate and conduct research studies. Several recent examples have appeared in the professional literature.

--Two important research developments in recent years are theories of career choice and development, and the development and use of computerized information systems. Data collection, processing, and analysis facilities are not, however, sufficiently developed in Idaho to take advantage of these developments in the foreseeable future.

CHAPTER VI

SYNTHESIS, SUMMARY, AND CONCLUSIONS

The reader who has covered all of the material in the preceding five chapters is to be commended for his dedication and persistence. Others, due to limited reading time, may have turned to this final chapter to get a quick overview of the contents of the report. The present chapter was written with both types of readers in mind. The reader who has digested the preceding material will find most of the main topics reflected or specifically discussed in the sections which follow. The reader who begins here, although missing several important details, hopefully will get a meaningful overview of the entire report. This chapter is divided into three main sections, entitled Synthesis, Summary, and Conclusions.

SYNTHESIS

The point of view stated at the beginning of this report was that appraisal instruments of all kinds were to be considered as means to accomplish certain objectives resulting from policy decisions. The writer also stated two basic premises which were intended to facilitate subsequent discussions and presentations. The first premise (see page 4) was that, if trained personnel and facilities are not available to do a job, it should not be attempted at all. The second premise (see pages 10-11) is more specifically related to the relationships observed between various predictors and vocational success criteria. The writer adopted a correlation of .45 as the minimum which should be obtained before one or more predictors could be considered suitable for guidance, selection, and placement.

To Test Or Not To Test

We might begin by stating it this way (with due apologies to Shakespeare): "To test or not to test, that is the question." This is not, however, the question at all, but rather one question (albeit an important one) that may be considered after other more important and basic questions are answered. Perhaps the most basic question of all is: "What provisions are we to make for the education and employment of all of our citizens?" As Munsterberg (1913, p. 18) pointed out, this kind of question lies outside the sphere of the applied sciences, and must be answered from ethical, philosophical, or religious bases. One example of how our democratic ideals have developed is given by Hand (1958, pp. 42-45), who states that they may be:

> . . . inferred from the historical documents which we consider great, the institutions which we support, the court decisions which we honor, the laws by which we govern ourselves, the customs to which we subscribe, and the qualities which we celebrate in our heroes and admire in our fellow citizens. (Hand, 1958, p. 43)

One ideal that may be derived from these basic sources is the supreme worth of each human life. If we subscribe to this ideal, vocational guidance of the best quality should follow naturally. Placement, in the sense of having the individual receive training and eventual employment as congruent as possible with his needs, abilities, and interests, also would be provided without question. Selection, on the other hand, would not be tolerated if it resulted in the exclusion of several individuals from educational training and employment opportunities.

The word "ideal" is appropriate in the foregoing discussion for obvious reasons. Whether we admit it or not, the supreme worth of the individual and equal rights and opportunities for all are not subscribed to by all members of our society. Moreover, local circumstances frequently prevent us, however idealistic we may be, from adequately serving the needs of each individual person. The important point is that only when more basic questions are considered and answered are we in a position to decide whether or not appraisal instruments are to be used. At one extreme, we may decide that, since tests are not perfect, we will dispense with them altogether. Needless to say, this does not constitute a satisfactory solution to several problems. At the other extreme we may decide to use tests, recognizing their imperfections, but be willing to tolerate the errors in judgment which result. This does not constitute a satisfactory solution either, but its justification lies in the fact that less errors result than in any other method available.

We might consider how decisions are to be made regarding the selection of persons for a work or training situation. The writer has heard certain individuals claim that they have a right to try any job they choose, and to attend any kind of institution or training class they wish. The desirability and possible outcomes of such an arrangement will be bypassed, as it appears unlikely that this kind of "individual liberty" will be realized to any great extent. A more realistic concern is how decisions are to be reached regarding which individuals will be selected for a work or training opportunity when applicants exceed available openings. This ignores, for purposes of discussion, actions which might be taken to bring the number of applicants and available openings into a better balance.

The reverse situation--more openings than applicants--obviously does not call for selection testing (disregarding certain obvious disqualifications, such as paralysis and gross mental deficiency), as practically everyone who applies is assured of acceptance. Implicit, however, is the question of whether the individual's choice is the best he could have made, and what kind--if any--of vocational guidance he received prior to arriving at this choice point. Once a group of individuals are chosen for work or training, individual differences and differing work or training assignments may require decisions regarding placement regardless of whether selection and vocational guidance have been operative. It is almost axiomatic that many groups of workers or trainees differ sufficiently so that different work assignments or different training methods are desirable. It is evident, therefore, that decision-making, based on individual appraisals, is potentially a factor in all three of the areas of vocational guidance, selection, and placement. At this point, having considered some rather basic questions and circumstances which require decision-making, "to test or not to test" emerges as yet another question which itself is subject to decision-making.

Appraisals Without Tests

If no individual appraisal whatsoever is performed in either vocational guidance or placement, it seems there is only one alternative. Even though stated in a different context, Holland's (1966, p. 91) statement about sums it up: ". . . the most efficient way to predict vocational choice is simply to ask the person what he wants to be . . . ." If this is all it amounts to, anyone can perform vocational guidance and placement (if, indeed, we can properly use the terms), and these functions have no status whatsoever as a professional activity.

Selection also degenerates into nothing more than a simple game of chance without appraisal, although matters become a little more complex as we consider this function. In simplest terms, a person selected for a work position or training course either will succeed or fail. It has been suggested that we flip a coin; if it turns up heads we select, and if tails we reject. A little reflection reveals, however, that it may not be this simple. We may, for example, have 10 times more applicants than we need. In this case a coin flip for each individual would not be very useful, unless the process was repeated a sufficient number of times to reduce the pool of applicants. Once we selected a group on the basis of some chance method like the flip of a coin, the best prediction for any single individual is the average of the group. Rate of failure, however, may be more important in a practical training or work situation. If the failure rate was about 16 out of 100, the flip of a coin might be less informative than the roll of a single die; by prior arrangement one face would mean rejection and the remaining five would mean acceptance. These schemes probably would not be too useful if the task at hand was to select 10 good applicants from, say, 100 who applied for work or training. If only 10 "bodies" were desired, they could be chosen by an even simpler method such as pointing a finger (and even this assumes some prior decision, however elementary). Thus, the placement function also loses any vestige of a professional activity.

All of this discussion thus far may have set the sophisticated reader to yawning with boredom, but the writer believes it is a necessary basis for the material to follow. Let us, therefore, continue, and enumerate several types of information, excluding formal test results, which enable us to make predictions better than chance concerning an individual's performance in a work or training situation.

Grade point average is probably the best single predictor of performance in further schooling, and is a good predictor in many future work situations.

Prior work history, as the individual gains experience, provides a basis for predictions of future work performance.

Years of schooling completed by the individual correlates substantially with intelligence (Anastasi, 1958, pp. 205-209), which in turn is a good predictor of many educational and vocational success criteria.

Socio-economic status as reflected by the occupation of the father indirectly provides evidence of academic and vocational potential.

Self estimates by students as they move from late junior high school into the senior high school years have predictive value, and perhaps should be encouraged and cultivated much more than we have been prone to do in the past. Holland's suggestion that we ask the individual what he wants to be may make more sense than seemed evident at first. Strong (1955, p. 186) cited studies of student estimates of their own test scores and grades which yielded correlations of .60 and higher. Rothney (1958, pp. 84-86) reported that, of 340 students given interpretations of test scores, about one-fourth seemed to feel that their expectations were confirmed, and more than half gave favorable responses. Less than 10 per cent gave a negative reaction.

The foregoing types of information, and others which might be developed, are by no means foolproof as bases for predictions. They do, however, remove us to some extent from the finger-pointing, coin-flipping and dice-rolling stages. The combination of self estimates by applicants reflected in their decisions to apply for various occupational training and work opportunities, and more or less formalized appraisal methods used by school officials and employment personnel result in a system that is workable. The world's work seems to get done and students are educated. A considerable amount of inefficiency and actual waste of human talent continues to exist, however, which prompts many responsible persons to consider still other means, including tests, to effect improvements.

Appraisals With Tests

It is rather surprising that, for a number of years, the belief persisted that better predictions could be made by intuition and other individual appraisal information than could be made with a battery of tests. The writer's experience has been that the "other information" referred to is frequently not made explicit, and thus may be more intuition than anything else. It is even more surprising that so few studies were completed which compared the accuracy of the two methods. Meehl (1954, p. 119) was able to locate only 16 to 20 studies of this type, and concluded that actuarial methods were approximately equal or superior to clinical methods in all but one study.

The writer, early in this report (see pages 3-4) took the position that tests, properly selected and used, can make a worthwhile contribution to program planning and related activities. Decision making is inescapable, and many decisions can be enhanced by the proper use of tests. Moreover, they can be compared with other appraisal information in local research studies, thus enabling the decision makers to sharpen and compare the different factors on which their decisions are based. Rather than thinking in terms of tests versus other predictive indexes, all types of information may be evaluated and the best chosen as bases for decision making. Tests should be rejected only after they have been given a fair chance to prove themselves in research studies, and in actual practice, not because of ignorance regarding their potential contributions.

The first premise stated by the writer, when applied to the use of tests, deserves some discussion. Tests vary tremendously in regard to the expertise required for administration and scoring. Problems surrounding the selection of tests for a given purpose, and their interpretation, are even more complex. Once a testing program is established, it still requires at least periodic professional supervision. Even carefully validated tests may have their predictive relationships altered by such factors as the following:

a. The mere passage of time may produce changes in the timeliness of test content, and the characteristics of the student population to which the tests are administered.

b. Extensive use of tests for selection may alter their predictive value. Grading systems are somewhat relative, and may change to fit the students selected. Tests validated on an unselected student group thus may lose some of their precision when they constitute the basis for selection.

c. Special treatments, such as remedial reading and special tutoring, may alter predictive relationships established before such special treatments were instituted.

d. New teaching personnel, new equipment and shop facilities, and new subject matter and media may affect correlations between tests and criteria established at an earlier date.

These are only a few examples which illustrate the need for continuous study and revision. This kind of responsibility should be specifically assigned to qualified professional personnel. If this is not done, the writer suggests that his first premise be given serious consideration: If trained personnel and facilities are not available to do a job, it should not be attempted at all.

The second premise regarding the minimum level of correlation necessary for practical applications of test results also merits a few final remarks. Early in this report (see pages 5-7) two basic functions of appraisal instruments were discussed, namely, predictions within groups and predictions of group membership. The correlation coefficient is useful for the first type of prediction. Since differential predictions are possible, correlations also have some utility for the second type of prediction. Predictions of group membership actually are less well developed, and have depended upon the degree of differentiation achieved by such instruments as interest inventories. The discriminant function does not provide a useful method for assigning students to more than two groups. Perhaps multivariate analysis of variance can be used as a basis for assigning students and workers to groups which they resemble most (Stewart, 1968, p. 171).

In either case, this writer's premise is that either predictions or group assignments should be equivalent to a minimum accuracy of 1 in 10 before the tests can be approved for practical use. This is an arbitrary choice, needless to say. Some readers may consider it too low, and possibly others will consider it higher than necessary. The writer would not argue the point; the crucial question in each case involving the precision with which tests predict a given criterion is: "Can you live with it?" The contents of the second and third chapters revealed that the 10 per cent level of accuracy could not be assumed for the present use of the General Aptitude Test Battery. On the other hand, the fourth and fifth chapters, as well as portions of the second and third chapters of this report, revealed that predictions exceeding an accuracy level of 10 per cent would be potentially possible. The necessary factor is that this capability be developed through the conduct of local research studies within the state of Idaho.

Outlook In Idaho

The study which culminated in the present report was conceived during 1966, and about this same time the writer began to lay the groundwork for later test validation studies. Officials of the Department of Employment, State of Idaho, were contacted, who agreed to administer all parts of the General Aptitude Test Battery to students entering the area vocational schools. This would provide a complete profile for each student on this test battery. During the interim, the Department of Employment assigned the responsibility of test administration and scoring to the area schools. The latter, due to limited staff time, have always had some difficulty in arranging for the collection of test information and other personal data for entering students. Consequently, complete test profiles have not been available for each student in all of the area vocational schools.

A statewide testing program, financed under the National Defense Education Act, was begun in the fall of 1961 in several Idaho secondary schools. The program consisted of the Differential Aptitude Tests (DAT), administered in the ninth grade, and the Iowa Tests of Educational Development (ITED), administered in the eleventh grade. Mr. Kenneth Armstrong, a doctoral student in guidance and counseling at the University of Idaho, presently is conducting a validation study under a subcontract with the State Occupational Research Unit. The study is a comparison of the GATB, the DAT, and the ITED as predictors of grade point averages of students who later graduated from the area vocational schools Mr. Armstrong also will use the results of the study as his doctoral thesis. The writer made arrangements with Mr. Elwyn DeLaurier, State Director of Guidance Services, to maintain a central file of results for the DAT and ITED so that they will be accessible for future research. As we go to press, the funding for this statewide testing program may be discontinued, thus eliminating or curtailing this source of useful test data.

Therefore, the outlook in Idaho for development of a statewide testing program for vocational students, and related research services, is not the best at this time. It may be possible to collect useful test data for students entering post-high school vocational programs from such sources as the American College Testing Program or the College Entrance Examination Board. A study directed by Dr. John L. Phillips, Jr., of Boise State College, Boise culminated about four years ago in a Counselor's Guide to Idaho Colleges (Phillips, 1965). Information was presented on Idaho colleges, and statistical information based on the ACT was included which made possible predictions of college grade point averages. The Counselor's Guide is due for revision soon, and the writer hopes to generate enough validation studies in the vocational schools of Idaho so that perhaps the next revision two or three years hence can contain vocational predictive data in addition to the present data on academic courses of study.

SUMMARY

This was a review and synthesis of research completed in the state of Idaho which pertained to the development and use of instruments for vocational guidance, selection, and placement. The first chapter set forth a point of view to the effect that tests and similar instruments were only means, and that their use (or non-use) must be determined by policy decisions. A basic premise was stated concerning the conduct of research studies and the use of appraisal instruments. If trained personnel are not available to carry out these functions, they should not be planned as part of an educational program or industrial personnel program. Vocational guidance, selection, and placement were briefly discussed and contrasted. Vocational guidance is oriented toward the individual, and is designed to assist him in a self appraisal, an understanding of the world of work, and then to select a career objective which maximizes his own satisfaction and his contributions to society. Selection and placement are more characteristic of the activities of educational institutions and business organizations which of necessity must decide who will be accepted and what roles or treatments will be expected of, or assigned to, those accepted. The view was expressed that existing appraisal instruments, properly researched and administered by trained professionals, can be used to advantage in all three functions of vocational guidance, selection, and placement.

The second and third chapters were concerned with research in which the General Aptitude Test Battery (GATB) was used to predict success in vocational training or work performance. Another basic premise was stated to the effect that predictions should be improved over chance methods so that the probability is equal to or better than 1 chance in 10, which is equivalent to a correlation coefficient of about .45 or higher. Since the GATB has had extensive use in Idaho, this battery and its accompanying research studies were examined carefully. Even though the GATB is perhaps the best battery of tests available for use in the vocational area, it was demonstrated that its use in Idaho cannot be approved without research studies in the local setting in which the battery is to be administered and interpreted. The third chapter contained the results of studies with the GATB as a predictor of success in vocational programs in Idaho area vocational schools. The results confirmed further the conclusions reached in the preceding chapter, and revealed that the published norms for different occupations were not always congruent with results obtained in the local setting. Moreover, the results indicated that other types of tests might be used in addition to, or in the place of, the GATB in an experimental prediction battery.

The fourth chapter presented a review of all research studies with appraisal instruments completed in Idaho, and the fifth chapter presented selected studies from elsewhere in the nation. Thirty-one studies were located and a summary of each was prepared for Chapter IV. More than three-fourths of the studies were completed as graduate research papers or theses. This resulted in research projects which reflected the individual interests of each researcher to the extent that the results generally were not comparable or additive from one study to another. Several of the studies were comparatively simple in design, and could have been improved in this respect. Several weaknesses were identified in such areas as sampling, statistical analyses, and the collection and treatment of data. The fifth chapter presented representative studies and reviews from elsewhere in the nation concerned with test validation, career choice and development theory, and computerized information systems. Research in the vocational area has received increased attention and prestige in recent years. Several of these national developments have important implications for future research in Idaho.

CONCLUSIONS

The research reported in the five preceding chapters, and citations from the research literature on appraisal instruments for vocational behavior, support the following conclusions:

--Tests and similar appraisal instruments, when administered and interpreted by trained professionals, can make a significant contribution to programs of vocational guidance, selection, and placement.

--Appraisal instruments should be evaluated for use by means of research studies in the local setting in which they are to be used.

--The General Aptitude Test Battery, even though it has been researched more extensively than perhaps any other test in the vocational area, can be recommended as only one tool in the kit of a trained counselor. Research completed to date is not a sufficient basis for its continued use in Idaho's program of vocational-technical education.

--A research program, staffed by professionally qualified personnel and adequately financed, should be developed in Idaho. Two activities which should receive high priority are (1) research on the predictive potential of various types of appraisal instruments, and (2) the establishment of policies and procedures regarding test administration which would result in complete and accurate test data.

--A wide variety of appraisal instruments has potential value as predictors of vocational success and membership in different vocational or training groups. An experimental battery should not consist only of aptitude tests, but also achievement, interest, and other types of appraisal instruments.

--A respectable number of research studies pertaining to vocational guidance, selection, and placement have been completed in Idaho. The studies generally reflect considerable research expertise, care in design and analysis, and adequate attention to tabular, graphic and written presentations. None of the studies completed, however, were entirely free from a number of weaknesses in design, analysis, and interpretation. All of the research completed to date does not constitute a satisfactory basis for the practical use of appraisal instruments in Idaho's system of vocational-technical education.

--Several important research projects elsewhere in the nation, and the research literature on vocational appraisal, form an important basis for continued efforts in Idaho. Principles pertaining to appraisal instruments and their use, and the types of instruments suitable for use in a variety of situations, are fairly well established. Relatively new research areas such as career choice and development theory, and computerized information systems, may need to await further developments in Idaho before they can be implemented on anything other than a limited research and experimental basis.

The reader who has covered all six chapters of this report will recall a number of instances in which the writer cited authors whose works were completed many decades ago. This was done on purpose to show that many useful ideas regarding the application of appraisal instruments have been with us for a long period of time. A quotation from Hull is appropriate as the present report draws to a close. Hull observed that the conception of specialized aptitudes is very ancient, and quoted from Book II of Plato's Republic a dialogue concerning the natural endowments of people most suited for the military profession.

> Following this, Plato proposes as a means of accomplishing the task, that persons being considered for the military profession shall be given "actions to perform" which shall test the retentiveness of their memories, their power of resistance to deceptions, of resistance to timidity and fear in terrifying situations, and to the seductions of pleasure. Thus we find Plato sketching forth very definitely a set of tests for military aptitude. Some twenty-three hundred years later the dream conceived by the Greek genius was realized in the United States army mental tests. Such is the halting course of progress. (Hull, 1928, p. 6)

It should not require another twenty-three hundred years for us to correct some of the deficiencies which continue to exist in the field of individual appraisal. We probably do not have that much additional time.

BIBLIOGRAPHY

Anastasi, Anne. Differential Psychology. Third edition, New York: The MacMillan Company, 1958.

Anastasi, Anne. Psychological Testing. Third edition, New York: The MacMillan Company, 1968.

Ayarra, Jr., David M., "An Investigation of the Relationship Between Self-Concept and Vocational Preference," Unpublished Masters Thesis, Idaho State University, Pocatello, Idaho, 1967.

Bennett, George K., et al. Manual For The Differential Aptitude Tests. Fourth edition, New York: The Psychological Corporation, 1966.

Berdie, Ralph F., "Aptitude, Achievement, Interest, and Personality Tests: a Longitudinal Comparison," Journal of Applied Psychology, Vol. 39, April, 1955, pp. 103-114.

Bertrand, Mary Jane, "An Experimental Study of the Measurement of Attitudes in Secondary School: Dropout Versus Stayin," Unpublished Masters Thesis, Idaho State University, Pocatello, Idaho, 1966.

Borow, Henry, "Modern Perspectives in Personnel Research." In N. B. Henry (Ed.), Personnel Services in Education. Fifty-eighth Yearbook of the National Society for the Study of Education, Part II, 1959, pp. 210-230.

Brenna, Lyle V., "A Study of the Occupational Plans of Senior Students From Representative Public Secondary Schools in Idaho," Unpublished Masters Thesis, College of Idaho, Caldwell, Idaho, 1961.

Brown, Merrill J., "A Survey of the Education and Employment Plans of Mountain Home High School's 1968 Graduates," Unpublished Seminar Paper Submitted for Education 290b, College of Idaho, Caldwell, Idaho, 1968.

Burke, Max E., "Predicting Performance in Industrial Electronics at North Idaho Junior College With The General Aptitude Test Battery," Unpublished Masters Thesis, University of Idaho, Moscow, Idaho, 1967.

Buros, Oscar Krisen (Ed.). Tests In Print. Highland Park, New Jersey: The Gryphon Press, 1961.

Buros, Oscar Krisen (Ed.). The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1959.

Buros, Oscar Krisen (Ed.). The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.

Cain II, E. Thomas, "Factors Operative in Curricular and/or Occupational Choice: A Study of Super's Theory," Unpublished Doctors Thesis, University of Idaho, Moscow, Idaho, 1969.

Cattell, Raymond B., "Some Deeper Significances of the Computer for the Practicing Psychologist," The Personnel and Guidance Journal, Vol. 44, October, 1965, pp. 160-166.

Chatburn, Acel H., "An Evaluation of the Program of Boise Junior College by Its Graduates," Unpublished Doctors Thesis, State College of Washington, Pullman, Washington, 1956.

Crites, John O., "Appraising The Appraisal Instruments," American Vocational Journal, Vol. 43, December, 1968, pp. 22-24ff.

Cronbach, Lee J. Essentials of Psychological Testing. Second edition, New York: Harper & Brothers, 1960.

Cronbach, Lee J. & Goldine C. Gleser. Psychological Tests and Personnel Decisions. Urbana, Illinois: University of Illinois Press, 1957.

Dorcus, Roy M. & Margaret Hubbard Jones. Handbook of Employee Selection. New York: McGraw-Hill Book Company, Inc., 1950.

Droege, Robert C., "Validity of USES Aptitude Test Batteries for Predicting MDTA Training Success," The Personnel and Guidance Journal, Vol. 46, June, 1968, pp. 984-989.

Faulks, Frederick G., "A Study of the Factors Influencing Recruitment of Students in the Agricultural Education Curriculum in Idaho," Unpublished Masters Thesis, University of Idaho, Moscow, Idaho, 1967.

Fifield, Marvin G. & Robert H. Jones. Early School Leavers in the Pocatello School District. Pocatello School District Number 25, Pocatello, Idaho, circa 1964.

Fifield, Marvin & Larry E. Watson. A Follow-up Study of Pocatello and Idaho Falls High School Graduates (1954-1963). Pocatello School District Number 25 and the State Occupational Research Unit, College of Education, University of Idaho, Moscow, Idaho, June, 1967.

Flanagan, John C. Project Talent Five-Year Follow-up (Preliminary Data). Palo Alto, California: American Institutes for Research, December, 1968.

Flanagan, John C., et al. Design for a Study of American Youth. Boston: Houghton Mifflin Company, 1962.

Flores, Thomas R., "A Comparison of Stability and Realism of Occupational Aspiration in Eighth-Grade and Twelfth-Grade Males," Unpublished Masters Thesis, Washington State University, Pullman, Washington, 1966.

Gardner, George F., "A Survey of Factors Influencing Farm Youth in Selecting College Curricula," Unpublished Masters Thesis, University of Idaho, Moscow, Idaho, 1957.

Garrett, Henry E. Statistics in Psychology and Education. Fourth edition, New York: Longmans, Green, and Company, 1953.

Gentry, N. Dale. A Follow-up Study of Vocational Students at North Idaho Junior College. Moscow, Idaho: College of Education, The State Occupational Research Unit, July, 1967.

Ghiselli, Edwin E., "Moderating Effects and Differential Reliability and Validity," Journal of Applied Psychology, Vol. 47, April, 1963, pp. 81-86.

Ghiselli, Edwin E. *The Measurement of Occupational Aptitude*. University of California Publications in Psychology, Vol. 8, No. 2. Berkeley, California: University of California Press, 1955.

Ghiselli, Edwin E. *The Validity of Commonly Employed Occupational Tests*. University of California Publications in Psychology, Vol. 5, No. 9. Berkeley, California: University of California Press, 1949.

Ghiselli, Edwin E. *The Validity of Occupational Aptitude Tests*. New York: John Wiley & Sons, Inc., 1966.

Hand, Harold C. *Principles of Public Secondary Education*. New York: Harcourt, Brace and Company, 1958.

Harris, Jo Ann, "The Computerization of Vocational Information," *The Vocational Guidance Quarterly*, Vol. 17, September, 1968, pp. 12-20.

Holland, John L. *The Psychology of Vocational Choice*. Waltham, Massachusetts: Blaisdell Publishing Company, 1966.

Holmquist, Larry P., "A Follow-up Study of the Future Farmers of America Members Who Earned the State Farmer Degree in Idaho From 1955 Through 1960," Unpublished Masters Thesis, University of Idaho, Moscow, Idaho, 1967.

Hull, Clark L., "An Automatic Machine for Making Multiple Aptitude Forecasts," *The Journal of Educational Psychology*, Vol. 16, December, 1925, pp. 593-598.

Hull, Clark L. *Aptitude Testing*. New York: World Book Company, 1928.

Impellitteri, Joseph T., "A Computerized Occupational Information System, *The Vocational Guidance Quarterly*, Vol. 15, June, 1967, pp. 262-264.

Jex, Frank B. *Predicting Academic Success Beyond High School*. Salt Lake City, Utah: University of Utah, 1966.

Johnston, Aileen, "A Descriptive Survey Indicating Some Curricular Changes in the Pocatello High Schools for the Non-College Bound Student With Implications for Job Availability," Unpublished Masters Thesis, Idaho State University, Pocatello, Idaho, 1967.

Kiler, Alvard R., "The Development of an Experimental Instrument to Ascertain Selected Individual Needs," Unpublished Masters Thesis, Idaho State University, Pocatello, Idaho, 1967.

Kuder, G. Frederic. *Kuder DD, Occupational Interest Survey*, General Manual. Chicago: Science Research Associates, Inc., 1966.

Lee, Dorothy, "An Occupational Follow-up Study of Students Who Were Graduated Since 1955 From The Secretarial Science Department of Boise Junior College Designed to Evaluate The Effectiveness of the Department," Unpublished Masters Thesis, College of Idaho, Caldwell, Idaho, 1958.

Loudermilk, Kenneth M., "A Review of the Literature Pertaining to Vocational Choice," Unpublished Term Paper, Psychology 210, University of Idaho, Moscow, Idaho, October, 1961. (Mimeographed)

Loudermilk, Kenneth M., "Prediction of Efficiency of Lumber and Paper Mill Employees," *Personnel Psychology*, Vol. 19, Autumn, 1966, pp. 301-310.

Loudermilk, Kenneth M., "The Relationship Between Aptitude, Personality, Physical Fitness, and Personal Data, and Job Performance in a Combined Lumber and Paper Mill," Unpublished Doctors Thesis, University of Idaho, Moscow, Idaho, 1964.

Loughary, John W., *et al.*, "Autocoun: A Computer-Based Automated Counseling Simulation System," *The Personnel and Guidance Journal*, Vol. 45, September, 1966, pp. 6-15.

McEwen, Gary N., "Relationships Between Training in Vocational Agriculture and Occupational Selection," Unpublished Masters Thesis, University of Idaho, Moscow, Idaho, 1962.

McHugh, William T., "A Discriminant Function Study of Factor Variables in Short Term Occupational Choice," Unpublished Masters Thesis, Idaho State University, Pocatello, Idaho, 1967.

McRae, Lorie J., "Follow Up Survey of Selected Graduates of the Idaho State University Trade and Technical School," Unpublished Masters Paper, Idaho State University, Pocatello, Idaho, 1968.

Meehl, Paul E. *Clinical Versus Statistical Prediction*. Minneapolis: University of Minnesota Press, 1954.

Meyerhoeffer, Gerald R., "The Relationship of the General Aptitude Test Battery, Reading Ability and Motivation to Success in Vocational Training," Unpublished Masters Paper, Washington State University, Pullman, Washington, 1967.

Michael, William B., "An Overview of the Symposium," *Educational and Psychological Measurement*, Vol. 18, Autumn, 1958, pp. 455-461.

Miller, John G. *Predictive Testing for Entrance in Vocational-Technical Schools*. New York University: Center for Field Research and School Services, August, 1968.

Minor, Frank J., *et al.*, "An Experimental Computer-Based Educational and Career Exploration System," *The Personnel and Guidance Journal*, Vol. 47, February, 1969, pp. 564-569.

Munsterberg, Hugo. *Psychology and Industrial Efficiency*. New York: Houghton Mifflin Company, 1913.

Nunnally, Jum C. *Psychometric Theory*. New York: McGraw-Hill Book Company, 1967.

Nybroten, Norman. *Labor Force and Career Plans of High School Seniors in Ten Counties of Southern Idaho*. Moscow, Idaho: Bureau of Business and Economic Research, University of Idaho, Monograph No. 4, December, 1964.

Olson, Cleon Y., "A Study of the Graduates of the Idaho Falls High School Class of 1959-60," Unpublished Masters Thesis, Brigham Young University, Provo, Utah, 1966.

Osipow, Samuel H. *Theories of Career Development*. New York: Appleton-Century-Crofts, 1968.

Parsons, Frank. *Choosing A Vocation*. New York: Houghton Mifflin Company, 1909.

Patterson, C.H., "Predicting Success in Trade and Vocational School Courses: Review of the Literature," *Educational and Psychological Measurement*, Vol. 16, Autumn, 1956, pp. 352-400.

Phillips, Jr., John L. (Ed.). *Counselor's Guide to Idaho Colleges*. Iowa City, Iowa: American College Testing Program, October, 1965.

Prediger, Dale J., *et al.*, "Predictors of Success in High School Level Vocational Education Programs: A Review, 1954-1967," *The Personnel and Guidance Journal*, Vol. 47, October, 1968, pp. 137-145.

Pucel, David J. & Howard F. Nelson. *General Aptitude Test Battery (B-1002 Form B) Training Success Norms*. University of Minnesota: Project MINI-Score, Department of Industrial Education, February, 1969.

Rothney, John W.M., *et al*. *Guidance Practices and Results*. New York: Harper & Brothers, 1958.

Rutledge, Gay G., "Survey of Dropouts in the Caldwell Junior and Senior High Schools for the School Year of 1964-65," Unpublished Masters Thesis, College of Idaho, Caldwell, Idaho, 1967.

Scharf, Jr., George C., "The Counselor Selection Scale: Design for an Experimental Test," Unpublished Doctors Thesis, University of Idaho, Moscow, Idaho, 1966.

Sorensen, Lloyd L., "The Occupational Status of Vocational Agriculture Graduates From Idaho High Schools Over the Past Ten Year Period 1956 Through 1965," Unpublished Masters Thesis, Utah State University, Logan, Utah, 1967.

State Occupational Research Unit. *A Survey of Literature Related to Selected Nonprofessional Occupations*. College of Education, University of Idaho, Moscow, Idaho, May, 1966.

Stewart, Lawrence H. *A Study of Certain Characteristics of Students and Graduates of Occupation-Centered Curricula*. Final Report, Project No. 5-0152, Contract No. OE-6-85-072, U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research. Berkeley, California: School of Education, University of California, June, 1968.

Strong, Edward K., Jr. *Vocational Interests 18 Years After College*. Minneapolis: University of Minnesota Press, 1955.

Strong, Edward K., Jr. *Vocational Interests of Men and Women*. Stanford University, California: Stanford University Press, 1943.

Super, Donald E. *Appraising Vocational Fitness*. New York: Harper & Brothers, 1949.

Super, Donald E., "The Multifactor Tests: Summing Up," The Personnel and Guidance Journal, Vol. 36, September, 1957, pp. 17-20.

Super, Donald E. & John O. Crites. Appraising Vocational Fitness. New York: Harper and Brothers, 1962. (Revised Edition)

Super, Donald E., et al. Career Development: Self-Concept Theory. New York: College Entrance Examination Board, 1963.

Super, Donald E., et al. The Vocational Maturity of Ninth-Grade Boys. Monograph Two. New York: Bureau of Publications, Teachers College, Columbia University, 1960.

Thorndike, Robert L. & Elizabeth Hagen. Ten Thousand Careers. New York: John Wiley and Sons, Inc., 1959.

Tiedeman, David V., et al. Information System For Vocational Decisions. Harvard Graduate School of Education, Annual Report 1966-1967, September, 1967.

Tiedeman, David V., et al. Information System For Vocational Decisions. Harvard Graduate School of Education, Second Annual Report 1967-1968, September, 1968.

Tiedeman, David V., et al., "The Multiple Discriminant Function--A Symposium," The Harvard Educational Review, Vol. 21, Spring, 1951, pp. 71-95.

U.S. Department of Labor. Manual for the General Aptitude Test Battery. Section III: Development. Washington, D.C.: U.S. Government Printing Office, October, 1967. (a)

U.S. Department of Labor. Summary of Occupational Validity on the General Aptitude Test Battery. USES Test Research Report No. 17, Bureau of Employment Security, November, 1967. (b)

van der Reis, A.P., "Is Retesting Justified in Personnel Selection?" Psychologia Africana, Vol. 10, January, 1963, pp. 19-30.

Watson, Larry E., "The Ten-Year Study--A Vertical Follow-up of Pocatello High School Graduates Over A Ten-Year Period (1954-1963)," Unpublished Masters Thesis, Idaho State University, Pocatello, Idaho, 1967.

Williamson, Charles F., "An Experimental Study to Test the Hypothesis That Interests May Be Determined by Reactions to Words With Muliple Definitions," Unpublished Masters Thesis, University of Idaho, Moscow, Idaho, 1964.

Witty, Bonita, "A Study of the Importance of English in Certain Occupations as Measured by College Requirements and Attitudes of Businessmen Toward English Competence," Unpublished Masters Thesis, College of Idaho, Caldwell, Idaho, 1962.

Young, James L., "Development of Qualification Tests For An Industrial Technology Electronics Program," Unpublished Masters Thesis, College of Idaho, Caldwell, Idaho, 1965.

APPENDIX A

F TESTS ON GATB SCORES FOR STUDENTS IN FOUR VOCATIONAL PROGRAMS AT BOISE STATE COLLEGE

This appendix contains four tables of data derived from analyses of the test performances of students enrolled in four vocational programs at Boise State College, Boise, Idaho. The four programs are as follows: auto mechanics, drafting and design, electronics, and welding. The contents of each table are arranged in identical form, and will be described briefly to assist the reader in interpretation.

Each table presents the results of F tests between the means of a group of graduates and a group of non-graduates. Each aptitude symbol in the first (left-hand) column is for a single factor score on the General Aptitude Test Battery (GATB), and these may be interpreted as follows:

G = General Learning Ability
V = Verbal Aptitude
N = Numerical Aptitude
S = Spatial Aptitude
P = Form Perception
Q = Clerical Perception
K = Motor Coordination
F = Finger Dexterity
M = Manual Dexterity

The second column (moving to the right) contains the number of cases. The upper number is for graduates, and the bottom number of each pair is for non-graduates. The third column contains the means of the score distributions, and their positions correspond with the number of cases in the second column. The fourth (right-hand) column contains the F ratios. A discussion of the contents of this appendix may be found on pages 31-32 of this report.

APPENDIX A (CONT'D)

Table 1

F Tests on GATB Scores for Graduates and Non-Graduates in Auto Mechanics

Aptitude[1]	N	Means	F Ratio*
G	53	104.5	.009
	31	104.7	
V	53	95.8	.117
	31	96.7	
N	53	101.6	.280
	31	100.3	
S	53	111.3	.018
	31	110.9	
P	53	108.5	.172
	31	107.3	
Q	53	100.6	.922
	31	98.2	
K	53	99.5	1.74
	31	103.9	
F	53	102.5	1.27
	31	98.2	
M	53	110.9	.432
	31	113.5	

*None of the F ratios are significant

[1] Please turn to page 99 for an interpretation of the aptitude symbols

APPENDIX A (CONT'D)

Table 2

F Tests on GATB Scores for Graduates and Non-Graduates in Drafting and Design

Aptitude[1]	N	Means	F Ratio*
G	62	113.3	.080
	20	114.1	
V	62	102.6	.547
	20	104.9	
N	62	106.3	.158
	20	105.3	
S	62	123.9	.839
	20	128.2	
P	62	115.4	.121
	20	114.2	
Q	62	103.8	.032
	20	104.3	
K	62	101.6	.078
	20	102.6	
F	62	105.2	2.85
	20	97.3	
M	62	110.3	.090
	20	111.6	

*None of the F ratios are significant

[1] Please turn to page 99 for an interpretation of the aptitude symbols

APPENDIX A (CONT'D)

Table 3

F Tests on GATB Scores for Graduates and Non-Graduates in Electronics

Aptitude[1]	N	Means	F Ratios*
G	20 22	114.0 110.2	1.05
V	20 22	101.9 101.0	.067
N	20 22	107.0 106.0	.104
S	20 22	122.9 119.0	.517
P	20 22	114.0 110.3	.949
Q	20 22	102.8 103.4	.041
K	20 22	98.3 103.1	.926
F	20 22	102.1 107.2	1.17
M	20 22	109.8 112.6	.269

*None of the F ratios are significant

[1] Please turn to page 99 for an interpretation of the aptitude symbols

APPENDIX A (CONT'D)

Table 4

F Tests on GATB Scores for Graduates and Non-Graduates in Welding

Aptitude[1]	N	Means	F Ratios
G	26	98.7	.320
	16	96.4	
V	26	91.9	.775
	16	95.2	
N	26	97.0	5.40*
	16	85.6	
S	26	106.1	.506
	16	109.4	
P	26	102.5	.041
	16	101.6	
Q	26	94.0	.251
	16	92.4	
K	26	93.2	.348
	16	89.4	
F	26	93.2	.905
	16	85.9	
M	26	95.9	.178
	16	99.8	

*Significant between the .05 and .01 levels

[1] Please turn to page 99 for an interpretation of the aptitude symbols

APPENDIX B

F TESTS ON GATB SCORES FOR STUDENTS IN SEVEN VOCATIONAL PROGRAMS AT IDAHO STATE UNIVERSITY

This appendix contains seven tables of data derived from analyses of the test performances of students enrolled in seven vocational programs at Idaho State University, Pocatello, Idaho. The seven programs are as follows: auto mechanics, business machine repair, cosmetology, drafting and design, electronics, instrumentation, and secretarial training. The contents of each table are arranged in identical form, and will be described briefly to assist the reader in interpretation.

Each table presents the results of F tests between the means of a group of graduates and a group of non-graduates. Each aptitude symbol in the first (left-hand) column is for a single factor score on the General Aptitude Test Battery (GATB), and these may be interpreted as follows:

G = General Learning Ability
V = Verbal Aptitude
N = Numerical Aptitude
S = Spatial Aptitude
P = Form Perception
Q = Clerical Perception
K = Motor Coordination
F = Finger Dexterity
M = Manual Dexterity

The second column (moving to the right) contains the number of cases. The upper number is for graduates, and the bottom number of each pair is for non-graduates. The third column contains the means of the score distributions, and their positions correspond with the number of cases in the second column. The fourth (right-hand) column contains the F ratios. A discussion of the contents of this appendix may be found on pages 35 (et passim) of this report.

APPENDIX B (CONT'D)

Table 1

F Tests on GATB Scores for Graduates and Non-Graduates in Auto Mechanics

Aptitude[1]	N	Means	F Ratios*
G	32	101.1	.201
	22	102.9	
V	32	92.3	.321
	22	94.0	
N	32	99.9	.442
	22	97.3	
S	32	108.5	1.63
	22	115.1	
P	32	105.5	.046
	22	104.6	
Q	32	99.1	1.58
	22	94.8	
K	32	102.9	1.20
	22	98.2	
F	32	98.1	.064
	20	96.8	
M	32	109.3	.744
	20	104.1	

*None of the F ratios are significant

[1] Please turn to page 104 for an interpretation of the aptitude symbols

APPENDIX B (CONT'D)

Table 2

F Tests on GATB Scores for Graduates and Non-Graduates in Business Machine Repair

Aptitude[1]	N	Means	F Ratios*
G	19	102.5	.001
	15	102.7	
V	19	95.3	1.76
	15	101.7	
N	19	100.5	.290
	15	98.6	
S	19	104.7	.202
	15	102.3	
P	19	105.7	.017
	15	105.2	
Q	19	101.3	.456
	15	98.4	
K	19	100.1	1.62
	15	106.9	
F	16	94.8	2.59
	15	105.8	
M	16	101.9	.815
	15	108.7	

*None of the F ratios are significant

[1] Please turn to page 104 for an interpretation of the aptitude symbols

Table 3

F Tests on GATB Scores for Graduates and Non-Graduates in Cosmetology

Aptitude[1]	N	Means	F Ratios
G	40 23	102.4 104.9	.580
V	40 23	99.7 104.3	1.93
N	40 23	99.2 95.9	.687
S	40 23	108.9 110.5	.134
P	40 23	114.6 116.5	.238
Q	40 23	112.0 112.6	.035
K	42 23	107.1 110.7	.816
F	28 15	115.5 102.3	4.67*
M	28 15	113.9 113.5	.005

*Significant between the .05 and .01 levels

[1] Please turn to page 104 for an interpretation of the aptitude symbols

Table 4

F Tests on GATB Scores for Graduates and Non-Graduates in Drafting and Design

Aptitude[1]	N	Means	F Ratios*
G	46 12	112.5 108.9	1.55
V	46 12	102.1 99.4	.654
N	46 12	105.7 100.9	2.87
S	46 12	124.9 119.8	1.21
P	46 12	115.0 113.2	.281
Q	43 8	104.5 104.1	.005
K	45 11	104.0 100.4	.477
F	44 9	99.5 106.9	.843
M	42 8	107.1 113.9	.631

*None of the F ratios are significant

[1] Please turn to page 104 for an interpretation of the aptitude symbols

Table 5

F Tests on GATB Scores for Graduates and Non-Graduates in Electronics

Aptitude[1]	N	Means	F Ratios
G	76	114.8	5.98*
	77	110.3	
V	76	103.8	1.17
	77	101.7	
N	76	109.6	.105
	77	109.0	
S	76	122.4	9.09**
	77	114.8	
P	76	110.9	.049
	77	110.4	
Q	71	104.5	1.06
	77	102.4	
K	70	98.4	1.68
	76	101.7	
F	57	101.5	.416
	55	99.2	
M	57	108.4	3.10
	56	101.5	

*Significant between the .05 and .01 levels
**Significant beyond the .01 level

[1] Please turn to page 104 for an interpretation of the aptitude symbols

Table 6

F Tests on GATB Scores for Graduates and Non-Graduates in Instrumentation

Aptitude[1]	N	Means	F Ratios
G	37 5	110.4 99.0	3.64
V	37 5	102.9 84.8	7.39**
N	37 5	103.2 101.4	.093
S	37 5	120.3 115.6	.319
P	37 5	113.6 105.8	1.22
Q	36 5	106.6 101.6	.697
K	34 4	100.8 104.8	.235
F	29 4	104.8 101.0	.141
M	27 3	110.6 107.0	.079

**Significant beyond the .01 level

[1] Please turn to page 104 for an interpretation of the aptitude symbols

Table 7

F Tests on GATB Scores for Graduates and Non-Graduates in Secretarial Training

Aptitude[1]	N	Means	F Ratios
G	183 67	106.2 104.1	1.34
V	183 67	103.7 104.9	.449
N	183 67	106.3 103.2	2.74
S	183 67	106.2 103.4	1.82
P	183 67	115.7 114.9	.127
Q	190 69	116.8 112.6	5.60*
K	190 69	113.1 113.4	.023
F	115 38	109.3 105.3	1.24
M	116 37	109.6 106.0	.810

*Significant between the .05 and .01 levels

[1] Please turn to page 104 for an interpretation of the aptitude symbols